

MIP Reformulations of the Probabilistic Set Covering Problem

Anureet Saxena ^{*†}

Vineet Goyal ^{*‡}

Miguel Lejeune ^{*§}

Abstract

In this paper we address the following probabilistic version (PSC) of the set covering problem: $\min\{cx \mid \mathbb{P}(Ax \geq \xi) \geq p, x_j \in \{0, 1\}^N\}$ where A is a 0-1 matrix, ξ is a random 0-1 vector and $p \in (0, 1]$ is the threshold probability level. We formulate (PSC) as a mixed integer non-linear program (MINLP) and linearize the resulting (MINLP) to obtain a MIP reformulation. We introduce the concepts of p -inefficiency and polarity cuts. While the former is aimed at reducing the number of constraints in our model, the later is used as a strengthening device to obtain stronger formulations. A hierarchy of relaxations for (PSC) is introduced, and fundamental relationships between the relaxations are established culminating with a MIP reformulation of (PSC) with no additional integer constrained variables. Simplifications of the MIP model which result when one of the following conditions hold are briefly discussed: A is a balanced matrix, A has the circular ones property, the components of ξ are pairwise independent, the distribution function of ξ is a stationary distribution or has the so-called disjunctive shattering property. We corroborate our theoretical findings by an extensive computational experiment on a test-bed consisting of almost 10,000 probabilistic instances. This test-bed was created using deterministic instances from the literature and consists of probabilistic variants of the set-covering model and capacitated versions of facility location, warehouse location and k-median models. Our computational results show that our procedure is orders of magnitude faster than any of the existing approaches to solve (PSC), and in many cases can reduce hours of computing time to fraction of seconds.

Keywords: Probabilistic Programming, Set Covering, Mixed Integer Programming, Cutting Planes.

^{*}Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, USA. Email: {anureets, vgoyal, mlejeune}@andrew.cmu.edu.

[†]Research was supported by the National Science Foundation through grant #DMI-0352885 and by the Office of Naval Research through contract N00014-03-1-0133.

[‡]Research supported in part by NSF grant CCF-0430751 and ITR grant CCR-0122581

[§]Corresponding Author

1 Introduction

In this paper we address the following probabilistic variant of the set-covering problem,

$$\begin{array}{ll} \min & cx \\ \text{s.t} & \mathbb{P}(Ax \geq \xi) \geq p \\ & x_j \in \{0, 1\} \quad j \in N \end{array} \quad (PSC)$$

where A is a 0-1 matrix defined on row-index set M and column-index set N , ξ is a 0-1 random M -vector, $p \in (0, 1]$ is the value of the *threshold probability* (also called the *reliability level*) and $c \in \mathbb{R}^N$ is the cost vector. Indeed, if we replace the probabilistic constraint $\mathbb{P}(Ax \geq \xi) \geq p$ in (PSC) by $Ax \geq 1$ we recover the well-known set covering problem.

(PSC) belongs to a class of optimization problems commonly referred to as probabilistic programs. Probabilistic programming was introduced by Charnes and Cooper [7] in the late fifties and has since then been studied extensively. We refer the reader to Prékopa [20] for a review of recent developments in this area. (PSC) is a very challenging problem in the field of stochastic mixed integer programming which combines inherent complexity of both mixed integer programming and stochastic programming. Several set-covering models which can be solved in a matter of seconds by state-of-art MIP solvers (such as CPLEX or XPRESS) can give rise to probabilistic problems which can take several minutes (at times hours) to solve [5]. One of the notions which has played a pivotal role in the algorithmic development of (PSC) is that of *p-efficiency*. Originally introduced by Prékopa [18], the concept of *p-efficiency* of a discrete probability distribution has been the focus of intense research in recent years (see [5, 6, 12, 16, 21]).

In a recent development, Beraldi and Ruszczyński [5] proposed an algorithm to solve (PSC). Their algorithm involves enumerating the complete set of *p-efficient* points of the distribution, and then solving a deterministic set covering problem for each one of the *p-efficient* points. Some discrete distributions can have an extremely large number of *p-efficient* points, even at a high reliability level, which makes the enumeration phase very expensive. Solving MIPs from each one of these *p-efficient* points is a different proposition altogether. Beraldi and Ruszczyński [5] experimented with some hybrid techniques to improve their algorithm, but concluded that the enumeration of *p-efficient* points continued to be the bottleneck in their procedure. The ongoing research of Luedtke et al. [17] considers problems very closely related to (PSC). The authors study stochastic problems containing joint chance constraints in which the probabilistic requirement is imposed on a linear inequality whose right-hand side is a random variable with finite support. They reformulate the stochastic problem as a mixed-integer programming problem that contains a number of 0/1 variables equal to the number of possible realizations of the random variable. More precisely, for each realization of the random variable, the authors introduce a mixing inequality and a knapsack constraint. A strengthened formulation of the MIP problem is then obtained by considering, for each realization, the set defined by the mixing inequality, and by deriving valid inequalities for

each mixing set. Clearly, the method proposed in [17] relies upon the concepts of mixing set and inequality, and differs from the solution approach of the present study.

In this paper, we propose a technique of encoding the enumeration phase in the algorithm of Beraldi and Ruszczyński [5], itself as a mixed integer program thereby combining (*MIPing*¹) the two phases into one integrated MIP. Expressing the enumeration problem as a mixed integer program allows us to use a state-of-art MIP solver to perform an *intelligent* enumeration thereby reaping the benefits of developments in the field of mixed-integer programming. Indeed, our computational experiments conducted over a test-bed of almost 10,000 probabilistic instances demonstrate that (PSC) derived from simple and moderately difficult set-covering problems can themselves be formulated as simple or moderately difficult MIPs. As a byproduct of our research, we introduce the concept of p -inefficiency and polarity cuts. While the former is aimed at reducing the number of constraints in our model, the latter is used as a strengthening device to obtain stronger formulations.

Following Beraldi and Ruszczyński [5], we do not make any assumption on the probability distribution of ξ , except that ξ can be decomposed into L blocks say $\{\xi^1, \dots, \xi^L\}$ such that ξ^t is a 0-1 random M_t -vector for $t \in \{1, \dots, L\}$ (where M_1, \dots, M_L is a partition of M), and ξ^i and ξ^j are independent random vectors for distinct i, j . Henceforth, for $z \in \mathbb{R}^M$ we denote by z^t the sub-vector of z formed by components in M_t for $t = 1 \dots L$. Furthermore, let $F : \{0, 1\}^M \rightarrow \mathbb{R}$ denote the cumulative distribution function of ξ and let F_t denote the restriction of F to M_t for $t = 1 \dots L$. In other words, for $z \in \{0, 1\}^M$, $F(z) = \mathbb{P}(\xi \leq z)$ and $F_t(z^t) = \mathbb{P}(\xi^t \leq z^t)$.

The rest of the paper is organized as follows. In Section 2 we formulate (PSC) as a mixed integer non-linear program (MINLP) and linearize it to obtain a MIP reformulation (MIP1). We introduce the concept of p -inefficiency and illustrate its application to derive a refined reformulation (MIP2) with fewer number of constraints. In Section 3 we discuss a class of cutting planes for (PSC), which we refer to as polarity cuts. We derive a linear programming based separation algorithm for these cuts and discuss techniques to reduce the coefficient matrix densities of the resulting linear programs. Section 4 examines the theoretical properties of polarity cuts and introduces a hierarchy of relaxations for (PSC). Fundamental relationships between the relaxations are established culminating with a MIP reformulation of (PSC) without introducing any more integer variables apart from the ones in the deterministic problem. We also discuss simplifications of the MIP reformulations which arise when A is a balanced matrix or has the circular ones property, or when all components of ξ are pairwise independent. Section 5 continues this discussion and addresses the case when F is a so-called stationary distribution. Finally, we discuss our computational results in Section 6. We corroborate our theoretical findings by a computational experiment conducted on a test-bed consisting of almost 10,000 probabilistic instances. This test-bed was constructed from probabilistic variants of set-covering models and capacitated versions of facility location, warehouse location and k-median models. In Section 7 we present some concluding remarks.

¹The phrase *MIPing* was coined by Matteo Fischetti and Andrea Lodi at the Ninth International meeting on Combinatorial Optimization (2005) at Aussois, France.

2 MIP Formulation

In this section we discuss a mixed integer programming formulation of (PSC). We need to express the probabilistic constraints $\mathbb{P}(Ax \geq \xi) \geq p$ by a system of linear inequalities involving continuous and integer constrained variables. In order to do so, we introduce auxiliary binary variables z_i for $i \in M$ and continuous variables η_t for $t \in \{1 \dots L\}$. We formulate the probabilistic constraints as choosing the right hand side z such that if $Ax \geq z$ then the covering constraint for the random vector is satisfied with probability at least p .

Theorem 2.1 (PSC) can be formulated as the following mixed integer non-linear program.

$$\begin{aligned}
 & \min_{(x,z,\eta)} && cx \\
 & \text{s.t.} && \\
 & Ax &\geq z \\
 & \sum_{t=1}^L \eta_t &\geq \ln p \\
 & \eta_t &\leq \ln F_t(z^t) \quad t \in \{1 \dots L\} \\
 & x_j &\in \{0, 1\} \quad \forall j \in N \\
 & z_i &\in \{0, 1\} \quad \forall i \in M
 \end{aligned} \tag{MINLP}$$

Proof: Suppose x is a feasible solution to (PSC). Let $z_i = \min(a_i^T x, 1)$ $i \in M$, where a_i denotes the i^{th} row of A ; $z_i = 1$ if and only if the i^{th} row of A is covered by x . Since, $p \leq \mathbb{P}(Ax \geq \xi) = \mathbb{P}(z \geq \xi) = \prod_{t=1}^L F_t(z^t)$, $\ln p \leq \sum_{t=1}^L \ln F_t(z^t)$. By letting $\eta_t = \ln F_t(z^t)$, we get a feasible solution (x, z, η) to (MINLP). Conversely, suppose (x, z, η) is a feasible solution to (MINLP). Since $\ln p \leq \sum_{t=1}^L \eta_t \leq \sum_{t=1}^L \ln F_t(z^t)$, it follows that $p \leq \prod_{t=1}^L F_t(z^t) = F(z) = \mathbb{P}(\xi \leq z) \leq \mathbb{P}(\xi \leq Ax)$ and x is a feasible solution to (PSC). ■

Next we discuss a linearization of MINLP which yields the desired mixed integer programming formulation.

Theorem 2.2 (MINLP) can be formulated as the following mixed integer program.

$$\begin{aligned}
 & \min_{(x,z,\eta)} && cx \\
 & \text{s.t.} && \\
 & Ax &\geq z && (MIP1) \\
 & \sum_{t=1}^L \eta_t &\geq \ln p \\
 & \eta_t &\leq (\ln F_t(v))(1 - \sum_{i \in M_t, v_i=0} z_i) \quad \forall v \in \{0, 1\}^{M_t} \quad \forall t \in \{1 \dots L\} \text{ s.t. } F_t(v) > 0 \\
 & 1 &\leq \sum_{i \in M_t, v_i=0} z_i \quad \forall v \in \{0, 1\}^{M_t} \quad \forall t \in \{1 \dots L\} \text{ s.t. } F_t(v) = 0 \\
 & x_j &\in \{0, 1\} \quad \forall j \in N \\
 & z_i &\in \{0, 1\} \quad \forall i \in M
 \end{aligned}$$

Proof: Suppose (x, z, η) is a feasible solution to (MINLP). By Theorem 2.1, x is a feasible solution to (PSC) and $\prod_{t=1}^L F_t(z^t) \geq p$, which implies that $F_t(z^t) > 0 \quad \forall t \in \{1 \dots L\}$. For $t \in \{1 \dots L\}$, let $v \in \{0, 1\}^{M_t}$. If $F_t(v) = 0$, then $z \not\leq v$ (since $F_t(z^t) > 0$) and $1 \leq$

$\sum_{i \in M_t, v_i=0} z_i$. Next consider the case when $F_t(v) > 0$. If $\sum_{i \in M_t, v_i=0} z_i \geq 1$ then $\eta_t \leq (\ln F_t(v))(1 - \sum_{i \in M_t, v_i=0} z_i)$ is trivially satisfied; if $\sum_{i \in M_t, v_i=0} z_i = 0$ then $z^t \leq v$ and $\eta_t \leq \ln F_t(z^t) \leq \ln F_t(v)$. Hence (x, z, η) is also a feasible solution to (MIP1). Conversely suppose (x, z, η) is a feasible solution to (MIP1). For $t \in \{1 \dots L\}$ and $v = z^t$, we have $F_t(v) > 0$ and $\eta_t \leq (\ln F_t(v))(1 - \sum_{i \in M_t, v_i=0} z_i) = \ln F_t(z^t)$ which implies that (x, z, η) is also a feasible solution to (MINLP). ■

Several comments are in order. First, note that (MIP1) has an exponential (2^{m_t}) number of constraints for each block $t = 1 \dots L$. This exponentiality in the number of constraints is unavoidable since the input specification for a *generic* distribution function F_t is itself exponential in the dimension m_t . Consequently, for generic distributions, (MIP1) has linear number of constraints in the input specification. Indeed, special properties of the distribution can be exploited to devise compact representations which can be integrated into the MIP model. We give examples of such structured distributions in Sections 4 and 5.

Second, (MIP1) remains a valid model even if only a subset of the set-covering constraints $Ax \geq 1$ are probabilistically constrained. Indeed in such a case, the variables z_i are defined only for those rows $i \in M$ which are probabilistically constrained and (MIP1) is amended to include the deterministic covering constraints. The same comment applies to the case when the set covering model has additional deterministic (not necessarily set-covering) constraints. This observation is significant since set-covering constraints often arise as sub-systems of more complex mixed integer programs; for instance models arising in manifold applications such as scheduling, service planning, location problem, logical data analysis, etc use set-covering constraints to model the coverage requirements.

Third, the constraint $\sum_{t=1}^L \eta_t \geq \ln p$ implies $\eta_t \geq \ln p$ (since $\eta_t \leq 0$) for $t = 1 \dots L$. If (x, z, η) is a feasible solution to (MIP1), $t \in \{1 \dots L\}$ and $v \in \{0, 1\}^{M_t}$, then $\ln p \leq \eta_t \leq (\ln F_t(v))(1 - \sum_{i \in M_t, v_i=0} z_i)$; if $F_t(v) < p$ then $\sum_{i \in M_t, v_i=0} z_i \geq 1$. Consequently, for $v \in \{0, 1\}^{M_t}$ satisfying $F_t(v) < p$, the corresponding constraint in (MIP1) can be replaced by the set-covering constraint $\sum_{i \in M_t, v_i=0} z_i \geq 1$. Furthermore, if $v, w \in \{0, 1\}^{M_t}$ such that $v \leq w$ and $F_t(w) < p$, then the set-covering constraint corresponding to v is dominated by the set-covering constraint arising from w . To summarize, we need to add the set-covering constraints only for *maximal* vectors $v \in \{0, 1\}^{M_t}$ which satisfy $F_t(v) < p$.

Definition 2.3 A point $v \in \{0, 1\}^m$ is called a *p-inefficient point* of the probability distribution function F if $F(v) < p$ and there is no binary point $w \geq v$, $w \neq v$ such that $F(w) < p$. The set of all *p-inefficient points* of F is called the *p-inefficient frontier* of F .

The notion of *p-inefficiency* is closely related to the notion of *p-efficiency* introduced by Prékopa [18]. A point $v \in \{0, 1\}^m$ is called a *p-efficient point* of the discrete probability distribution function F if $F(v) \geq p$ and there is no binary point $w \leq v$, $w \neq v$ such that $F(w) \geq p$. The set of all *p-efficient points* of F is called the *p-efficient frontier* of F . If S is the set of binary vectors which are either *p-efficient* or dominate a *p-efficient* point, and T is the set of binary vectors which are either *p-inefficient* or are dominated by a *p-inefficient* point, then $\{S, T\}$ defines a partition of the lattice $\{0, 1\}^m$.

The following theorem summarizes the above discussion. For $t \in \{1 \dots L\}$, let S_t denote the set of binary vectors which are either p -efficient or dominate a p -efficient point of F_t and let I_t denotes the set of p -inefficient points of F_t .

Theorem 2.4 *(MIP1) can be reformulated as the following mixed integer program.*

$$\begin{aligned}
& \min_{(x,z,\eta)} && cx \\
& s.t && \\
& Ax &\geq z \\
& \sum_{t=1}^L \eta_t &\geq \ln p \\
& \eta_t &\leq (\ln F_t(v))(1 - \sum_{i \in M_t, v_i=0} z_i) & \forall v \in S_t \ \forall t \in \{1 \dots L\} \\
& 1 &\leq \sum_{i \in M_t, v_i=0} z_i & \forall v \in I_t \ \forall t \in \{1 \dots L\} \\
& x_j &\in \{0, 1\} \ \forall j \in N \\
& z_i &\in \{0, 1\} \ \forall i \in M
\end{aligned} \tag{MIP2}$$

The key difference between (MIP1) and (MIP2) is that, unlike (MIP1), (MIP2) has constraints arising only from points which are either p -inefficient or dominate a p -inefficient point of F_t . In theory, the number of such points can be exponential; however, in practice the number of such points is only a small fraction of the total number of lattice points.

Figure 2 shows this information graphically for two distributions, namely Circular and Star (see [5] for the definition of these distributions). The horizontal axis represents the reliability level p , whereas the vertical axis gives the number of points which are either p -inefficient or dominate a p -inefficient point, averaged over 1000 randomly generated instantiations (block size = 10) of the Circular and Star distributions. Note that, even at a reliability level of 0.8 less than 20% of the lattice points qualified the above condition. As an interesting consequence of this observation, it follows that our model can handle significantly large block sizes provided there exists an efficient algorithm to enumerate points which lie on or above the the p -inefficient frontier. Such an enumeration algorithm, clearly, would need to exploit properties of the specific distribution involved and its investigation goes beyond the scope of this paper.

Let (MIP2') denote the relaxation of (MIP2) obtained by replacing the integrality conditions on z variables by $0 \leq z \leq 1$. Let $\text{opt}(\text{MIP2})$ and $\text{opt}(\text{MIP2}')$ denote the optimal values of (MIP2) and (MIP2'), respectively. The proposition that follows shows that (MIP2') is also a valid reformulation of (PSC).

Proposition 2.5 $\text{opt}(\text{MIP2}) = \text{opt}(\text{MIP2}')$.

Proof: It suffices to show that $\text{opt}(\text{MIP2}) \leq \text{opt}(\text{MIP2}')$. Suppose (x, z, η) is an optimal solution to (MIP2'). Let $\bar{z} \in \{0, 1\}^M$ be defined as, $\bar{z}_i = 1$ if and only if $a_i^T x \geq 1$ where a_i denotes the i^{th} row of A for $i \in M$. Note that for $v \in S_t \cup I_t$ ($t = 1 \dots L$), $\sum_{i \in M_t, v_i=0} z_i \leq \sum_{i \in M_t, v_i=0} \bar{z}_i$, which implies that (x, \bar{z}, η) is also a feasible solution to (MIP2), and $\text{opt}(\text{MIP2}) \leq \text{opt}(\text{MIP2}')$. ■

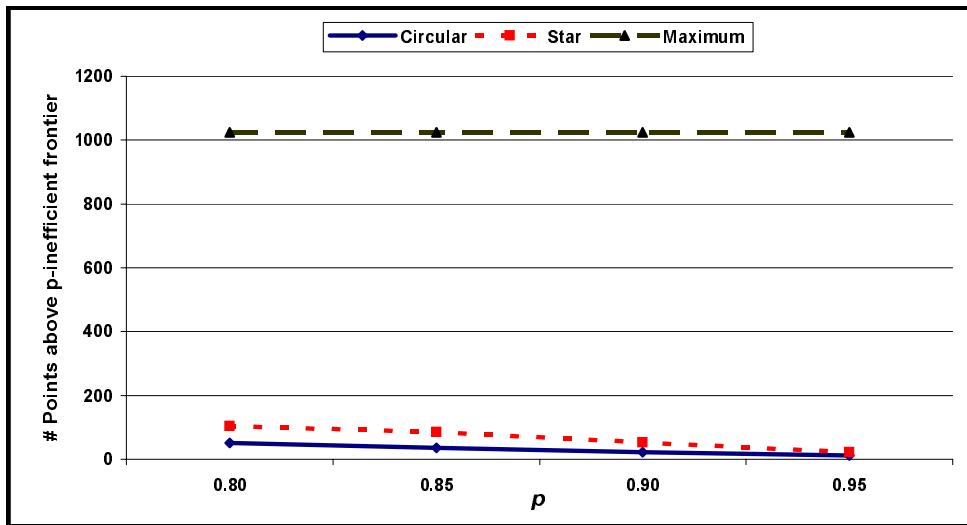


Figure 1: Number of points above the p -inefficient frontier

(MIP2) and (MIP2') provide two contrasting alternatives to solve (PSC). While (MIP2') has fewer number of binary variables, the integrality constraints on the z variables in (MIP2) can be used to generate strong valid cutting planes which might assist the overall solution procedure. We used a hybrid model in our experiments which combines the attractive features of both of these models. In Section 3 we discuss a class of strong valid cutting planes for the (MIP2) formulation, called polarity cuts, which do not cut off the optimal solution to (MIP2'). In our computational experiments we strengthened the (MIP2') formulation at the root node by these polarity cuts, and then applied CPLEX to the strengthened formulation. As our computational results (Section 6) demonstrate, this hybrid model is substantially better than (MIP2) and (MIP2'), considered individually.

We conclude this section by a small example. Consider the following probabilistic set-covering problem in which the reliability level p is equal to 0.8 and the right hand side is a 5-dimensional random 0-1 vector whose cumulative distribution function is given in Table 2.

$$\begin{aligned}
 & \min \sum_{j=1}^5 x_j \\
 & \text{s.t.} \\
 & \mathcal{P} \left(\begin{array}{l} x_1 + x_2 \geq \xi_1 \\ x_2 + x_3 \geq \xi_2 \\ x_3 + x_4 \geq \xi_3 \\ x_4 + x_5 \geq \xi_4 \\ x_5 + x_1 \geq \xi_5 \end{array} \right) \geq 0.8 \\
 & x_j \in \{0, 1\} \quad j = 1, \dots, 5
 \end{aligned} \tag{2.1}$$

0-1 Vector v	$\mathbb{P}(\xi \leq v)$	$\ln \mathbb{P}(\xi \leq v)$	Type of v	0-1 Vector v	$\mathbb{P}(\xi \leq v)$	$\ln \mathbb{P}(\xi \leq v)$	Type of v
00000	0.36834	-0.99876	T	10000	0.36834	-0.99876	T
00001	0.36834	-0.99876	T	10001	0.37243	-0.98769	T
00010	0.36834	-0.99876	T	10010	0.36834	-0.99876	T
00011	0.37243	-0.98769	T	10011	0.37658	-0.97663	T
00100	0.36834	-0.99876	T	10100	0.36834	-0.99876	T
00101	0.36834	-0.99876	T	10101	0.37243	-0.98769	T
00110	0.95672	-0.04424	S	10110	0.95672	-0.04424	S
00111	0.96736	-0.03318	S	10111	0.97812	-0.02212	S
01000	0.36834	-0.99876	T	11000	0.37243	-0.98769	T
01001	0.36834	-0.99876	T	11001	0.37658	-0.97663	T
01010	0.36834	-0.99876	T	11010	0.37243	-0.98769	T
01011	0.37243	-0.98769	T	11011	0.38077	-0.96557	I
01100	0.37243	-0.98769	T	11100	0.37658	-0.97663	T
01101	0.37243	-0.98769	T	11101	0.38077	-0.96557	I
01110	0.96736	-0.03318	S	11110	0.97812	-0.02212	S
01111	0.97812	-0.02212	S	11111	1.00000	0.00000	S

Table 1: Cumulative Probability Distribution

The first column of the table contains a 5-dimensional binary vector v , the second column gives the value of $\mathbb{P}(\xi \leq v)$ while the third column contains $\ln \mathbb{P}(\xi \leq v)$. The fourth column categorizes the binary vector v into one of the following three categories; (S): v is either p -efficient or dominate a p -efficient point, (I): v is p -inefficient and (T): v is dominated by a p -inefficient point. The distribution represented by table 2 has 8 points of type S and 2 points of type I . The (MIP2) formulation for this probabilistic instance is given by,

$$\begin{aligned}
& \min \sum_{j=1}^5 x_j \\
& \text{s.t} \\
& \quad x_1 + x_2 \geq 1 \\
& \quad x_2 + x_3 \geq 1 \\
& \quad x_3 + x_4 \geq 1 \\
& \quad x_4 + x_5 \geq 1 \\
& \quad x_5 + x_1 \geq 1 \\
& \quad \eta_1 \geq -0.22314 \quad (\ln 0.8 = -0.22314) \\
& \quad \eta_1 \leq -0.04424(1 - z_1 - z_2 - z_5) \\
& \quad \eta_1 \leq -0.03318(1 - z_1 - z_2) \\
& \quad \eta_1 \leq -0.03318(1 - z_1 - z_5) \\
& \quad \eta_1 \leq -0.04424(1 - z_2 - z_5) \\
& \quad \eta_1 \leq -0.02212(1 - z_1) \\
& \quad \eta_1 \leq -0.02212(1 - z_2) \\
& \quad \eta_1 \leq -0.02212(1 - z_5) \\
& \quad \eta_1 \leq 0 \\
& \quad z_3 \geq 1 \\
& \quad z_4 \geq 1 \\
& \quad x_j \in \{0, 1\} \quad j = 1 \dots 5 \\
& \quad z_i \in \{0, 1\} \quad i = 1 \dots 5
\end{aligned} \tag{2.2}$$

3 Polarity Cuts

For the sake of brevity, we assume in this section that (PSC) has only one block ($L = 1$). The results discussed here can be easily extended to the more general case ($L \geq 1$) by applying them to each one of the blocks independently. Recall that I denotes the set of p -inefficient points of F and S denotes the set of 0-1 points which are either p -efficient or dominate a p -efficient point of F . Consider the following set of constraints which constitute (MIP2).

$$\eta \leq (\ln F(v))(1 - \sum_{i \in M, v_i=0} z_i) \quad \forall v \in S \quad (3.3)$$

$$\sum_{i \in M, v_i=0} z_i \geq 1 \quad \forall v \in I \quad (3.4)$$

$$z_i \in \{0, 1\} \quad \forall i \in M \quad (3.5)$$

These constraints define the so-called big-M formulation of $P = \text{clconv}\{(z, \eta) \mid z \in \{0, 1\}^M, \eta \leq \ln F(z), F(z) \geq p\}$ (clconv stands for closed convex hull). A central question in polyhedral analysis is to examine the strength of the defining inequalities (3.3), (3.4) with respect to the underlying integer hull P .

For the sake of discussion, consider the example (2.1) introduced in section 2. We generated the following complete minimal description of P using the PORTA [8] software.

$$\begin{aligned} \eta &\leq -0.04424 + 0.02212z_1 + 0.01106z_2 + 0.01106z_5 \\ \eta &\leq -0.04424 + 0.01106z_1 + 0.02212z_2 + 0.01106z_5 \\ \eta &\leq -0.04424 + 0.01106z_1 + 0.01106z_1 + 0.02212z_5 \\ \eta &\leq -0.04424 + 0.02212z_2 + 0.02212z_5 \\ z_3 &= 1 \\ z_4 &= 1 \\ 0 &\leq z_1 \leq 1 \\ 0 &\leq z_2 \leq 1 \\ 0 &\leq z_5 \leq 1 \end{aligned}$$

Each one of the above inequalities defines a facet of P . Note that the constraints in the MIP2 formulation (2.2) derived from points in S do not define facets of P ; in fact, they do not even define non-empty faces of P . Indeed the constraint $\eta_1 \leq -0.04424(1 - z_1 - z_2 - z_5)$ (derived from $(0, 0, 1, 1, 0, -0.04424) \in P$) is strictly dominated by the facet-defining inequality $\eta \leq -0.04424 + 0.02212z_1 + 0.01106z_2 + 0.01106z_5$. This suggests that the inequalities derived from points in S can be significantly strengthened by coefficient tightening procedures. Strengthening inequalities by coefficient tightening has two shortcomings. First, such procedures are sequence dependent and produce different inequalities depending on the order in which the coefficients are examined. Consequently, several different inequalities can be obtained by strengthening a single inequality, and it is difficult to decide a priori which of these will strengthen the formulation most effectively. Second, such procedures can generate

only a subset of all valid (facet-defining) inequalities of P . Next we describe a procedure to generate valid inequalities of P which overcomes both of these shortcomings.

The lemma that follows provides crucial insights into the polyhedral structure of P . Let $J = \{i \in M \mid z_i = 1 \ \forall z \in \{0, 1\}^M \text{ s.t. } F(z) \geq p\}$. Let $e \in \{1\}^M$ denote a vector of ones and e^i ($i \in M$) denote the i^{th} unit vector.

Lemma 3.1 *For $i \in J$, $e - e^i$ is a p -inefficient point of F . Furthermore, $\dim(P) = m + 1 - |J|$ and for $i \in M \setminus J$, $z_i \leq 1$ defines a facet of P . If $\alpha z - \beta \eta \geq \Delta$ defines a facet of P different from the ones defined by $z_i \leq 1$ ($i \in M \setminus J$), then $\beta \geq 0$, $\alpha_i \geq 0 \ \forall i \in M \setminus J$ and $\sum_{i \in M \setminus J} \alpha_i + \beta > 0$.*

Proof: Clearly for $i \in J$, $F(e - e^i) < p$ and $e - e^i$ is a p -inefficient point of P . Furthermore, since $P \subseteq \{(z, \eta) \mid z_i = 1 \ \forall i \in J\}$, $\dim(P) \leq m + 1 - |J|$. To see that $\dim(P) = m + 1 - |J|$, consider the following $m + 2 - |J|$ affinely independent points in P , $\{(e - e^i, \ln F(e - e^i)) \mid i \in M \setminus J\} \cup \{(e, 0), (e, -1)\}$. Using a similar construction, it can be shown that $z_i \leq 1$ defines a facet of P for $i \in M \setminus J$. Suppose $\alpha z - \beta \eta \geq \Delta$ defines a facet of P different from the ones defined by $z_i \leq 1$ ($i \in M \setminus J$). Since P recedes in the direction $(z = 0, \eta = -1)$, $\beta \geq 0$. For $i \in M \setminus J$, there exists $z \in \{0, 1\}^M$ and $\eta \in \mathbb{R}$ such that $(z, \eta) \in P$, $\alpha z - \beta \eta = \Delta$ and $z_i = 0$; since $(z + e^i, \eta) \in P$, $\alpha z + \alpha_i - \beta \eta \geq \Delta$ which implies that $\alpha_i \geq 0 \ \forall i \in M \setminus J$. ■

Theorem 3.2 *Let $(\hat{z}, \hat{\eta}) \in \mathbb{R}^M \times \mathbb{R}$ such that $0 \leq \hat{z} \leq 1$ and $\sum_{i \in M, v_i=0} \hat{z}_i \geq 1 \ \forall v \in I$. $(\hat{z}, \hat{\eta}) \in P$ if and only if the optimal value of the following linear program is non-negative.*

$$\begin{aligned}
& \min_{(\alpha, \beta, \Delta)} && \alpha \hat{z} - \beta \hat{\eta} - \Delta \\
& \text{s.t.} && \\
& && \alpha z - \beta \ln(F(z)) - \Delta \geq 0 \quad z \in S \\
& && \sum_{i \in M \setminus J} \alpha_i + \beta = 1 \\
& && \alpha_i \geq 0 \quad i \in M \setminus J \\
& && \alpha_i = 0 \quad i \in J \\
& && \beta \geq 0
\end{aligned} \tag{3.6}$$

Furthermore, if (α, β, Δ) is a feasible solution to (3.6) satisfying $\alpha \hat{z} - \beta \hat{\eta} - \Delta < 0$, then $\alpha z - \beta \eta \geq \Delta$ is a valid inequality for P which cuts off $(\hat{z}, \hat{\eta})$.

Proof: Clearly, if (α, β, Δ) is a feasible solution to (3.6), then $\alpha z - \beta \eta \geq \Delta$ is a valid inequality for P ; hence if $(\hat{z}, \hat{\eta}) \in P$ then the optimal value of (3.6) is non-negative. Conversely, suppose $(\hat{z}, \hat{\eta}) \notin P$. For $i \in J$, $e - e^i$ is a p -inefficient point of F (Lemma 3.1) and hence $\hat{z}_i = 1$. Consequently, there exists a facet defining inequality $\alpha z - \beta \eta \geq \Delta$ of P which cuts off $(\hat{z}, \hat{\eta})$. Without loss of generality, we can assume that $\alpha_i = 0 \ \forall i \in J$. Since $0 \leq \hat{z} \leq 1$, the facet defined by $\alpha z - \beta \eta \geq \Delta$ is different from the ones defined by $z_i \leq 1 \ i \in M \setminus J$, which implies that $\alpha_i \geq 0 \ \forall i \in M \setminus J$, $\beta \geq 0$ and $\sum_{i \in M \setminus J} \alpha_i + \beta > 0$ (Lemma 3.1). If $\theta = \sum_{i \in M \setminus J} \alpha_i + \beta$, then $\frac{1}{\theta}(\alpha, \beta, \Delta)$ is feasible solution of P , $\frac{1}{\theta}(\alpha \hat{z} - \beta \hat{\eta} - \Delta) < 0$ and the optimal value of (3.6) is negative. ■

Several comments are in order. First, the above theorem yields a systematic procedure for iteratively strengthening the (MIP2) formulation by generating cutting planes which cut

off the incumbent fractional solution in each iteration. Furthermore, unlike the coefficient tightening procedure, the above separation procedure is guaranteed to produce every valid (facet-defining) inequality of P . Cuts derived using the separation linear program (3.6) are referred to as *polarity cuts* in the sequel.

Second, the linear program (3.6) has lot more constraints than the number of variables, which suggests that the dual simplex algorithm is the most suitable linear programming algorithm for solving (3.6); the associated basis is a $(m+1) \times (m+1)$ matrix. Thus for $m = 10$, the dual simplex method updates the inverse of a 11×11 basis matrix. Third, we introduce a penalty term $\sum_{i \in M} w \alpha_i$ in the objective function, where $w = 10^{-4}$, which is aimed at favoring sparse cuts over equally good dense cuts (see Fischetti and Lodi [14] and de Souza and Balas [22] for importance of sparse cuts in cutting plane procedures).

Fourth, note that $(\hat{z}, \hat{\eta}) \in P$ if and only if $(e - \hat{z}, \hat{\eta}) \in \bar{P}$ where $\bar{P} = \{(z, \eta) \mid (e - z, \eta) \in P\}$ and e is a vector of ones. In other words, we can apply an affine transformation $(z, \eta) \mapsto (e - z, \eta)$ to $(\hat{z}, \hat{\eta})$, solve the separation linear program (3.6) in the transformed space and apply the inverse transformation to the cut (if any). The advantage of such a transformation is the reduction in the number of non-zeros in the coefficient matrix of (3.6) thereby improving the overall performance of the dual simplex algorithm due to sparsity considerations. To see this, note that 0-1 points $z \in S$ have significantly more number of ones than zeros. Figure 2 illustrates this phenomenon for Circular and Star distributions graphically. The horizontal axis represents the threshold probability p . The vertical axis represents the ratio of the number of non-zeros in the separation linear program formulated in the transformed space and original space, respectively, averaged over 1000 randomly generated instantiations of each one of the distributions (block size = 10). As is evident from Figure 2, the above transformation can reduce the density of the coefficient matrix by 30 – 60%.

We conclude this Section by revisiting example (2.1) introduced in Section 2. The separation linear program for this example in the original and transformed space is given by,

$$\begin{array}{c}
 \left[\begin{array}{l}
 \min \quad \sum_{i=1}^5 \alpha_i \hat{z}_i - \beta \hat{\eta} - \Delta \\
 s.t. \quad \alpha_3 + \alpha_4 + 0.04424\beta - \Delta \geq 0 \\
 \alpha_3 + \alpha_4 + \alpha_5 + 0.03318\beta - \Delta \geq 0 \\
 \alpha_2 + \alpha_3 + \alpha_4 + 0.03318\beta - \Delta \geq 0 \\
 \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + 0.02212\beta - \Delta \geq 0 \\
 \alpha_1 + \alpha_3 + \alpha_4 + 0.04424\beta - \Delta \geq 0 \\
 \alpha_1 + \alpha_3 + \alpha_4 + \alpha_5 + 0.02212\beta - \Delta \geq 0 \\
 \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + 0.02212\beta - \Delta \geq 0 \\
 \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 - \Delta \geq 0 \\
 \alpha_1 + \alpha_2 + \alpha_5 + \beta = 1 \\
 \alpha_1, \alpha_2, \alpha_5, \beta \geq 0 \\
 \alpha_3, \alpha_4 = 0
 \end{array} \right] \quad \text{Original Space}
 \end{array}
 \quad
 \begin{array}{c}
 \left[\begin{array}{l}
 \min \quad \sum_{i=1}^5 -\alpha_i(1 - \hat{z}_i) - \beta \hat{\eta} - \Delta \\
 s.t. \quad -\alpha_1 - \alpha_2 - \alpha_5 + 0.04424\beta - \Delta \geq 0 \\
 -\alpha_1 - \alpha_2 + 0.03318\beta - \Delta \geq 0 \\
 -\alpha_1 - \alpha_5 + 0.03318\beta - \Delta \geq 0 \\
 -\alpha_1 + 0.02212\beta - \Delta \geq 0 \\
 -\alpha_2 - \alpha_5 + 0.04424\beta - \Delta \geq 0 \\
 -\alpha_2 + 0.02212\beta - \Delta \geq 0 \\
 -\alpha_5 + 0.02212\beta - \Delta \geq 0 \\
 -\Delta \geq 0 \\
 \alpha_1 + \alpha_2 + \alpha_5 + \beta = 1 \\
 \alpha_1, \alpha_2, \alpha_5, \beta \geq 0 \\
 \alpha_3, \alpha_4 = 0
 \end{array} \right] \quad \text{Transformed Space}
 \end{array}$$

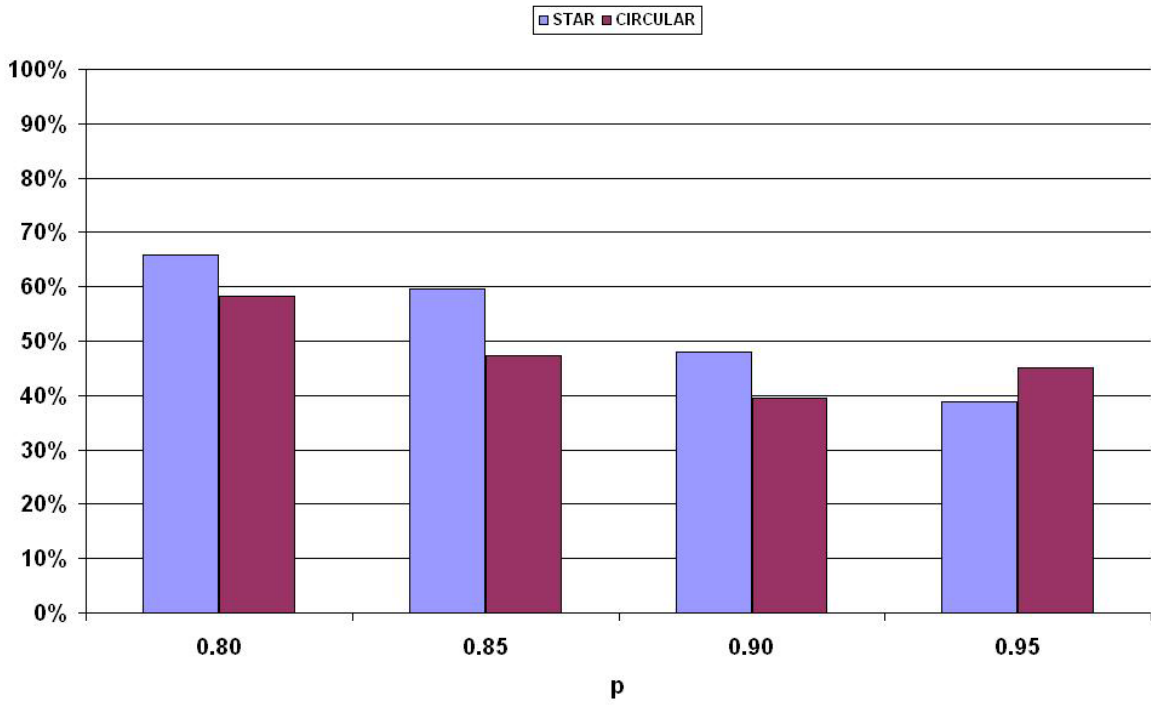


Figure 2: Reduction in density of the coefficient matrix of Separation Linear Program (3.6)

4 Hierarchy of Relaxations

In this section, we use the concept of polarity cuts developed in Section 3 to introduce a hierarchy of relaxations for (PSC). We establish fundamental relations between the relative strengths of these relaxations and conclude with a strengthened reformulation of (MIP2) which gets rid of the integrality constraints on the z variables at the cost of introducing additional linear constraints.

For $t = 1 \dots L$, let

$$P_t = \text{clconv} \{ (v, \eta) \mid v \in S_t, \eta \leq \ln(F_t(v)) \}$$

and let

$$P = \text{clconv} \left\{ (z, \eta) \left| \begin{array}{l} z \in \{0, 1\}^M, \eta \in \mathbb{R}^L \\ z^t \in S_t \forall t = 1 \dots L \\ \eta_t \leq \ln(F_t(z^t)) \forall t = 1 \dots L \\ \sum_{t=1}^L \eta_t \geq \ln(p) \end{array} \right. \right\}.$$

Let (R1) denote the LP relaxation of (MIP2) obtained by replacing the integrality constraints on x and z variables by $0 \leq x \leq 1$ and $0 \leq z \leq 1$, respectively. Similarly, let (R5) denote the relaxation of (MIP2) obtained by replacing the integrality constraints on x variables by

$0 \leq x \leq 1$. Let (R2)-(R4) and (R6) be defined as,

$$\min_{(x,z,\eta)} \left\{ cx \mid \begin{array}{l} Ax \geq z, 0 \leq x, z \leq 1 \\ \sum_{t=1}^L \eta_t \geq \ln(p) \\ (z^t, \eta_t) \in P_t \forall t = 1 \dots L \end{array} \right\} \quad (R2)$$

$$\min_{(x,z,\eta)} \left\{ cx \mid \begin{array}{l} Ax \geq z, 0 \leq x, z \leq 1 \\ \eta \in \mathbb{R}^L \\ (z, \eta) \in P \end{array} \right\} \quad (R3)$$

$$\min_{(x,z,\eta)} \left\{ cx \mid \begin{array}{l} Ax \geq z, 0 \leq x, z \leq 1 \\ \eta \in \mathbb{R}^L \\ (z, \eta) \in P \\ z \in \{0, 1\}^M \end{array} \right\} \quad (R4)$$

$$\min_{(x,z,\eta)} \left\{ cx \mid \begin{array}{l} Ax \geq z, 0 \leq x, z \leq 1 \\ \sum_{t=1}^L \eta_t \geq \ln(p) \\ (z^t, \eta_t) \in P_t \forall t = 1 \dots L \\ x \in \{0, 1\}^N \end{array} \right\} \quad (R6)$$

Let $\text{opt}(R1)$ denote the optimal value of (R1); $\text{opt}(R2)$, $\text{opt}(R3)$, $\text{opt}(R4)$, $\text{opt}(R5)$, $\text{opt}(R6)$ and $\text{opt}(MIP2)$ are defined similarly.

Theorem 4.1 *(R1-R6) are relaxations of (MIP2), and*

$$\text{opt}(R1) \leq \text{opt}(R2) \leq \text{opt}(R3) \leq \text{opt}(R4) = \text{opt}(R5) \leq \text{opt}(R6) = \text{opt}(MIP2) \quad (4.7)$$

Proof: Clearly, (R1), (R2), (R5) and (R6) are relaxations of (MIP2); if (x, z, η) is a feasible solution to (MIP2) then $(z, \eta) \in P$ and $z \in \{0, 1\}^M$ which implies that (R3) and (R4) are also relaxations of (MIP2). The only non-trivial statements in (4.7) are $\text{opt}(R2) \leq \text{opt}(R3)$ and $\text{opt}(R6) = \text{opt}(MIP2)$, which we prove next.

Note that,

$$P = \text{clconv} \left\{ (z, \eta) \mid \begin{array}{l} z \in \{0, 1\}^M, \eta \in \mathbb{R}^L \\ z^t \in S_t \forall t = 1 \dots L \\ \eta_t \leq \ln(F_t(z^t)) \forall t = 1 \dots L \\ \sum_{t=1}^L \eta_t \geq \ln(p) \end{array} \right\} \quad (4.8)$$

$$\subseteq \text{clconv} \left\{ (z, \eta) \mid \begin{array}{l} z \in \{0, 1\}^M, \eta \in \mathbb{R}^L \\ z^t \in S_t \forall t = 1 \dots L \\ \eta_t \leq \ln(F_t(z^t)) \forall t = 1 \dots L \end{array} \right\} \cap \left\{ (z, \eta) \mid \sum_{t=1}^L \eta_t \geq \ln(p) \right\} \quad (4.9)$$

$$= \left\{ (z, \eta) \mid (z^t, \eta_t) \in P_t \forall t = 1 \dots L \right\} \cap \left\{ (z, \eta) \mid \sum_{t=1}^L \eta_t \geq \ln(p) \right\} \quad (4.10)$$

where (4.9) follows from $clconv(\Pi_1 \cap \Pi_2) \subseteq clconv(\Pi_1) \cap clconv(\Pi_2)$ for any two sets Π_1 and Π_2 . (4.8–4.10) show that every feasible solution to (R3) is also a feasible solution to (R2) and hence $opt(R2) \leq opt(R3)$.

Finally to show that $opt(R6) = opt(MIP2)$, it suffices to show that $opt(R6) \geq opt(MIP2)$. Let (x, z, η) be an optimal solution to (R6). Let $\bar{z} \in \{0, 1\}^M$ be defined as, $\bar{z}_i = 1$ if and only if $a_i^T x \geq 1$ where a_i denotes the i^{th} row of A for $i \in M$. Let $\bar{\eta}_t = \ln F_t(\bar{z}^t)$ for $t = 1 \dots L$.

Let $t \in \{1 \dots L\}$. Since $(z^t, \eta_t) \in P_t$, there exist $(v^{tj}, \eta_{tj}) \in P_t$ and $0 < \lambda_{tj} \leq 1$ for $j = 1 \dots n_t$ (for some $n_t \geq 1$) such that $z^t = \sum_{j=1}^{n_t} \lambda_{tj} v^{tj}$, $\eta_t \leq \sum_{j=1}^{n_t} \lambda_{tj} \eta_{tj}$, $\sum_{j=1}^{n_t} \lambda_{tj} = 1$, and $v^{tj} \in \{0, 1\}^{M_t}$. Note that $v^{tj} \leq \bar{z}^t$ and $\eta_{tj} \leq \bar{\eta}_t$ for $j = 1 \dots n_t$. Furthermore, there exists $j \in \{1 \dots n_t\}$ such that $\eta_t \leq \eta_{tj}$ which implies that $\eta_t \leq \bar{\eta}_t$.

Hence $\eta_t \leq \bar{\eta}_t$ for $t = 1 \dots L$, $\sum_{t=1}^L \bar{\eta}_t \geq \sum_{t=1}^L \eta_t \geq \ln(p)$, $(x, \bar{z}, \bar{\eta})$ is a feasible solution to (MIP2) and $opt(R6) \geq opt(MIP2)$. \blacksquare

Several comments are in order. First and foremost, note that (R6) is a valid reformulation of (MIP2). In fact, the proof of the above theorem gives a simple linear time algorithm to construct an optimal solution to (MIP2) from an optimal solution to (R6). Second, the constraint $(z^t, \eta_t) \in P_t$ in (R6) can be replaced by a system of inequalities which define P_t ; these inequalities, in turn, can be separated efficiently using the separation algorithm discussed in Section 3.

Third, while (R6) and (MIP2') are both valid reformulations of (MIP2) containing no additional integer constrained variables, the LP relaxation of (R6) (namely R2) is much stronger than the LP relaxation of (MIP2'), as confirmed by our computational results (see Section 6). Besides, special properties of the distribution can at times be used to represent the condition $(z^t, \eta_t) \in P_t$ in (R6) compactly using a polynomial number of additional constraints and variables, yielding a formulation which can be used to address large scale problems. See Section 5 for an example of such a distribution.

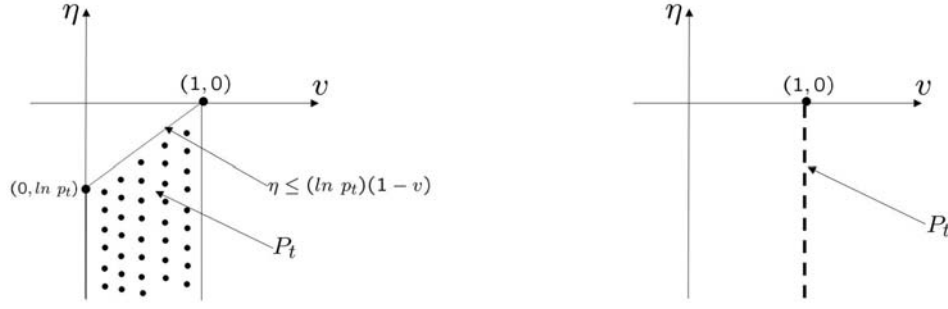
Fourth, (R2) and (R3) can be regarded as two extremes of a series of relaxations of (PSC). To see this, suppose that $\Gamma_1 \dots \Gamma_k$ is a partition of $\{1 \dots L\}$. Consider the following optimization problem,

$$\min \left\{ cx \mid \begin{array}{ll} Ax & \geq z, \ 0 \leq x, z \leq 1 \\ \sum_{t=1}^L \eta_t & \geq \ln(p) \\ (z(\Gamma_j), \eta(\Gamma_j)) & \in P(\Gamma_j) \ j = 1 \dots k \end{array} \right\} \quad (R(\Gamma))$$

where for $j = 1 \dots k$,

$$P(\Gamma_j) = clconv \left\{ (v, \eta) \mid \begin{array}{ll} v & \in \{0, 1\}^{\cup_{t \in \Gamma_j} M_t}, \ \eta \in \mathbb{R}^{\Gamma_j} \\ \eta_t & \leq \ln \mathbb{P}(\xi_i \leq v_i \ \forall i \in M_t) \ \forall t \in \Gamma_j \\ \mathbb{P}(\xi_i \leq v_i \ \forall i \in M_t) & \geq p \ \forall t \in \Gamma_j \\ \sum_{t \in \Gamma_j} \eta_t & \geq \ln(p) \end{array} \right\},$$

$z(\Gamma_j)$ is the sub-vector of z formed by components in $\cup_{t \in \Gamma_j} M_t$ and $\eta(\Gamma_j) \in \mathbb{R}^{\Gamma_j}$ such that the t^{th} component of $\eta(\Gamma_j)$ is equal to η_t for $t \in \Gamma_j$. Indeed, if all Γ_j are singletons then $R(\Gamma)$



$$p_t \geq p$$

$$p_t < p$$

Figure 3: P_t for the case when $|M_t| = 1$

reduces to (R2), whereas if $k = 1$ then $R(\Gamma)$ reduces to (R3). Note that for any partition $\Gamma_1 \dots \Gamma_k$ of $\{1 \dots L\}$, $R(\Gamma)$ is a relaxation of (MIP2), and

$$\text{opt}(R2) \leq \text{opt}(R(\Gamma)) \leq \text{opt}(R3) .$$

In other words, among all relaxations $R(\Gamma)$, (R2) is the weakest relaxation while (R3) is the strongest one.

Fifth, from a computational standpoint, (R1) and (R2) can be solved as linear programs, whereas (R4), (R5) and (R6) can be solved as mixed integer programs. To our best knowledge any method to solve (R3) will involve generating a complete description of P or enumerating all the extreme points and extreme rays of P .

Sixth, consider the case when all components of the random 0-1 vector ξ are pairwise independent. In other words, $L = m$ and $|M_t| = 1$ for $t = 1 \dots L$. If $M_t = \{i\}$ and $p_t = \mathbb{P}(\xi_i \leq 0)$ then

$$P_t = \begin{cases} \{(v, \eta) \mid 0 \leq v \leq 1, \eta \leq (\ln p_t)(1 - v)\} & \text{if } p_t \geq p \\ \{(v, \eta) \mid v = 1, \eta \leq 0\} & \text{if } p_t < p \end{cases}$$

for $t = 1 \dots L$ (see Figure 3). Consequently, in this case the (R6) relaxation is identical to (MIP2').

A fundamental question in polyhedral combinatorics is to determine necessary and sufficient conditions under which optimal value of the relaxation \mathcal{R} of a combinatorial optimization problem \mathcal{P} coincides with the optimal value of \mathcal{P} . For instance, it is well known that the LP relaxation of a pure integer program has integer optimal solution for every choice of the cost vector and right hand side, if and only if the coefficient matrix of the integer program is Totally Unimodular. The proposition that follows gives a similar result about (PSC). Recall that a 0-1 matrix A is *ideal* if the polytope $\{x \mid Ax \geq 1, 0 \leq x \leq 1\}$ is integral. Similarly, a 0-1 matrix A is *balanced* if every submatrix of A is ideal (Theorem 6.1 of Cornuejols [11], also see Conforti and Cornuejols [10]).

Proposition 4.2 *Given a 0-1 matrix A , the following two statements are equivalent.*

1. $\text{opt}(R5) = \text{opt}(MIP2)$ for every cost vector c , cumulative distribution function F and threshold probability p .
2. A is a balanced matrix.

Proof: Please refer to the Appendix. ■

In some special cases, structural properties of the matrix A can be used to design MIP reformulations of (PSC) with fewer number of binary variables as compared to (MIP2). The proposition that follows gives example of one such property. Recall that a 0-1 matrix A has the circular ones property if in every row either the ones or zeros are consecutive. Set covering models involving 0-1 matrices with circular ones property often arise in scheduling problems [3].

Proposition 4.3 *If A has the circular ones property, then (MIP2) can be reformulated as, ($n = |N|$)*

$$\begin{aligned}
 & \min_{(x,z,\eta,y)} && cx \\
 & s.t && \\
 & Ax &\geq z \\
 & \sum_{t=1}^L \eta_t &\geq \ln p \\
 & \eta_t &\leq (\ln F_t(v))(1 - \sum_{i \in M_t, v_i=0} z_i) \quad \forall v \in S_t \quad \forall t \in \{1 \dots L\} \\
 & 1 &\leq \sum_{i \in M_t, v_i=0} z_i \quad \forall v \in I_t \quad \forall t \in \{1 \dots L\} \\
 & \sum_{j \in N} x_j &= y \\
 & x_j &\leq 1 \quad j \in N \\
 & x_j &\geq 0 \quad j \in N \\
 & z_i &\in \{0, 1\} \quad \forall i \in M \\
 & y &\in \{0, 1, \dots, n\}
 \end{aligned} \tag{4.11}$$

Proof: Please refer to the Appendix. ■

We conclude this section by revisiting example (2.1) introduced in Section 2. Note that in this example $L = 1$ and the complete description of P_1 was given in Section 3. The reformulation (R6) of (2.1) is given by,

$$\begin{aligned}
\min \quad & \sum_{j=1}^5 x_j \\
\text{s.t.} \quad & x_1 + x_2 \geq 1 \\
& x_2 + x_3 \geq 1 \\
& x_3 + x_4 \geq 1 \\
& x_4 + x_5 \geq 1 \\
& x_5 + x_1 \geq 1 \\
& \eta_1 \geq -0.22314 & (\ln 0.8 = -0.22314) \\
& \eta_1 \leq -0.04424 + 0.02212z_1 + 0.01106z_2 + 0.01106z_5 \\
& \eta_1 \leq -0.04424 + 0.01106z_1 + 0.02212z_2 + 0.01106z_5 \\
& \eta_1 \leq -0.04424 + 0.01106z_1 + 0.01106z_1 + 0.02212z_5 \\
& \eta_1 \leq -0.04424 + 0.02212z_2 + 0.02212z_5 \\
& \eta_1 \leq 0 \\
& z_3 = 1 \\
& z_4 = 1 \\
& z_i \leq 1 \quad i = 1 \dots 5 \\
& z_i \geq 0 \quad i = 1 \dots 5 \\
& x_j \in \{0, 1\} \quad j = 1 \dots 5
\end{aligned} \tag{4.12}$$

Note that the coefficient matrix of (2.1) has the circular ones property. Consequently, an alternative reformulation of (MIP2) is obtained by replacing the last three constraints of (4.12) by,

$$\begin{aligned}
\sum_{j=1}^5 x_j &= y \\
z_i &\in \{0, 1\} \quad i = 1 \dots 5 \\
y &\in \{0, 1, \dots, 5\} \\
x_j &\leq 1 \quad j = 1 \dots 5 \\
x_j &\geq 0 \quad j = 1 \dots 5.
\end{aligned}$$

5 Stationary Distributions

In the previous section, we demonstrated how structural properties of the matrix A can be used to devise reformulations of (MIP2) with fewer number of integer variables. In this section, we investigate the same question apropos of the probability distribution function F .

Our model (MIP2) is meant for generic distributions and makes no assumption on the random vector ξ , except that ξ can be expressed as a cartesian product of groups of mutually independent random vectors. The practical applicability of our model is restrained by the size of input specification which is exponential in the case of generic distributions. While the full generality of our model is useful, specially structured distributions can often lead to models which can be used to address large scale problems. Next, we give an example of such a specially structured distribution.

A cumulative distribution function $F : \{0, 1\}^M \rightarrow \mathbb{R}$ is said to be *stationary* if $F(v) = F(w) \forall v, w \in \{0, 1\}^M$ such that $\sum_{i \in M} v_i = \sum_{i \in M} w_i$. Thus the value $F(z)$ of a stationary distribution depends only on the number ones in z . More precisely, any stationary distribution is completely defined by a vector $(\lambda_0, \dots, \lambda_m)$ ($m = |M|$) where λ_i represents the value of the stationary distribution at a lattice point with exactly i ones. The proposition that follows gives a polynomial (in m) sized MIP reformulation of (MIP2) for the case of stationary distributions.

Proposition 5.1 Suppose for $t \in \{1 \dots L\}$, F_t is a stationary distribution defined by the vector $(p_0^t, p_1^t \dots p_{m_t}^t)$ ($m_t = |M_t|$) and $k_t = \min \{k \mid 0 \leq k \leq m_t, p_k^t \geq p\}$. (MIP2) is equivalent to,

$$\begin{array}{ll}
\min_{(x,z,\eta,y,w)} & cx \\
s.t & \\
& Ax \geq z \\
& x_j \in \{0,1\} \quad j \in N \\
& z_i \in \{0,1\} \quad i \in M \\
& \sum_{t=1}^L \eta_t \geq \ln p \\
& \left. \begin{array}{l} y_t = \sum_{i \in M_t} z_i \\ y_t \geq k_t \\ y_t \in \mathbb{Z} \end{array} \right\} t = 1 \dots L \\
& \left. \begin{array}{l} y_t + w_k^t(n-k) \leq n \\ y_t + w_k^t(k+1) \geq k+1 \\ \eta_t \leq (\ln p_k^t) w_k^t \\ w_k^t \in \{0,1\} \end{array} \right\} k = k_t \dots m_t, \quad t = 1 \dots L
\end{array} \tag{MIP4}$$

Proof: Suppose (x, z, η, y, w) is a feasible solution to (MIP4). Let $t \in \{1 \dots L\}$. Note that $y_t = \sum_{i \in M_t} z_i$ is equal to the number of ones in z^t . Furthermore, $w_k^t = 1$ if and only if $y_t \leq k$ for $k = k_t \dots m_t$. The constraint $y_t \geq k_t$ ensures that $\mathbb{P}(\xi_i \leq z_i \mid i \in M_t) \geq p$ and $z^t \in S_t$, whereas the constraint $\eta_t \leq (\ln p_k^t) w_k^t$ ensures that $\eta_t \leq \ln \mathbb{P}(\xi_i \leq z_i \mid i \in M_t)$. Hence (x, z, η) is a feasible solution to (MIP2). Conversely, suppose (x, z, η) is a feasible solution to (MIP2); (x, z, η, y, w) is a feasible solution to (MIP4) where y, w are defined as: for $t = 1 \dots L$, $y_t = \sum_{i \in M_t} z_i$ and for $k = k_t \dots m_t$, $w_k^t = 1$ if and only if $y_t \leq k$. ■

Note that (MIP4) has linear (in m) number of additional variables and constraints, and hence can be used to handle arbitrarily large block sizes. Suppose $F : \{0,1\}^M \rightarrow \mathbb{R}$ is a stationary cumulative distribution function defined by the vector $(p_0 \dots p_m)$. It is worth observing that the lattice $\{0,1\}^M$ associated with F can be partitioned into $(m+1)$ slices such that the k^{th} slice is composed of 0-1 M -vectors with exactly k ones for $k = 0, 1 \dots m$, and the closed convex hull of the set $\{(z, \eta) \mid z \in \{0,1\}^M, \sum_{i \in M} z_i = k, \eta \leq \ln(F(z))\}$ has a compact description given by $\{(z, \eta) \mid 0 \leq z \leq 1, \sum_{i \in M} z_i = k, \eta \leq \ln p_k\}$. In other words, stationary distributions possess the disjunctive shattering property defined below.

Definition 5.2 A cumulative distribution function $F : \{0,1\}^M \rightarrow \mathbb{R}$ is said to possess the disjunctive shattering property (DSP) if the lattice $\{0,1\}^M$ can be partitioned into polynomial (in m) number of subsets, say $\{0,1\}^M = \cup_{j=1}^k M(j)$, such that the closed convex hull of the set $\{(z, \eta) \mid z \in M(j), \eta \leq \ln(F(z))\}$ has a polynomial (in m) sized compact description $\{(z, \eta) \mid A^j z + d^j \eta \geq b^j\}$ for $j = 1 \dots k$.

The proposition that follows gives a polynomial sized reformulation of (MIP1) for the case when each one of the distribution functions F_t possess the disjunctive shattering property.

Proposition 5.3 For $t \in \{1 \dots L\}$, suppose F_t possesses the disjunctive shattering property and the lattice $\{0, 1\}^{M_t}$ corresponding to F_t is partitioned into k_t subsets, say $\{0, 1\}^{M_t} = \cup_{j=1}^{k_t} M_t(j)$, and for $j = 1 \dots k_t$, $clconv \{(v, \eta) \mid v \in M_t(j), \eta \leq \ln F_t(v)\} = \{(v, \eta) \mid A^{jt}v + d^{jt}\eta \geq b^{jt}\}$. (MIP1) is equivalent to,

$$\begin{aligned}
& \min_{(x, z, \eta, \tilde{z}, \tilde{\eta}, \lambda)} \quad cx \\
& s.t. \quad \left. \begin{aligned}
Ax &\geq z \\
x_j &\in \{0, 1\} \quad j \in N \\
\sum_{t=1}^L \eta_t &\geq \ln p \\
z_i &= \sum_{j=1}^{k_t} \tilde{z}_i^{jt} \quad i \in M_t \\
\eta_t &= \sum_{j=1}^{k_t} \tilde{\eta}^{jt} \\
\sum_{j=1}^{k_t} \lambda_{jt} &= 1 \\
A^{jt}\tilde{z}^{jt} + d^{jt}\tilde{\eta}^{jt} &\geq \lambda_{jt}b^{jt} \quad j = 1 \dots k_t \\
\lambda_{jt} &\geq 0 \quad j = 1 \dots k_t
\end{aligned} \right\} t = 1 \dots L
\end{aligned} \tag{MIP5}$$

Proof: Let $t \in \{1 \dots L\}$. Note that $P_{jt} = \{(v, \eta) \mid A^{jt}v + d^{jt}\eta \geq b^{jt}\} \neq \emptyset \quad \forall j = 1 \dots k_t$, and the following constraints define the extended formulation (see Balas [1]) of $P_t = clconv(\cup_{j=1}^{k_t} P_{jt})$ for $t = 1 \dots L$.

$$\begin{aligned}
z_i &= \sum_{j=1}^{k_t} \tilde{z}_i^{jt} \quad i \in M_t \\
\eta_t &= \sum_{j=1}^{k_t} \tilde{\eta}^{jt} \\
\sum_{j=1}^{k_t} \lambda_{jt} &= 1 \\
A^{jt}\tilde{z}^{jt} + d^{jt}\tilde{\eta}^{jt} &\geq \lambda_{jt}b^{jt} \quad j = 1 \dots k_t \\
\lambda_{jt} &\geq 0 \quad j = 1 \dots k_t
\end{aligned}$$

Consequently, (MIP4) is equivalent to (R6) and the above proposition follows immediately from Theorem 4.1. \blacksquare

Note that (MIP5) has polynomial (in m) number of additional variables and constraints, and has no additional integer constrained variables. (MIP5) can be generalized to handle the case when some or all of the system of inequalities $A^{jt}z^{jt} + d^{jt}\eta^{jt} \geq b^{jt}$ have exponential number of inequalities, provided there exists a polynomial time separation algorithm to identify a violated inequality among $A^{jt}z^{jt} + d^{jt}\eta^{jt} \geq b^{jt}$. The generalization, however, is technical and is of limited interest in the context of the current paper.

6 Computational Results

Figure 4 gives the flowchart of our algorithm to solve (PSC). We implemented our algorithm using COIN-OR [9] and CPLEX (version 9.0). The linear programming module (OsiClp) of COIN-OR was used to solve all resulting linear programs, while the final MIP formulation

Deterministic Problem	Source	Total Number of Instances	Number of Instances retained	Instances Excluded	# Constraints	# Columns	# Probabilistic Constraints
Set Covering	ORLIB [4]	80	60	scplcr 10-13 scpcyc 06-11 scpnrg 1-5 scpnrh 1-5	50-500	500-5000	50-500
SSCFLP	Holmberg [15]	71	70	p58	60-230	510-6030	50-200
CWLP	ORLIB [4] (Set 1)	37	37	-	66-100	816-2550	50
k-median	ORLIB [4]	20	20	-	101-201	2550-10100	50-100

Table 2: Test Bed of Deterministic Instances

was solved using CPLEX 9.0. All experiments were carried out on a 2GHz P4 processor with 2GB RAM. In this section, we describe the computational results of our experiment on a test-bed consisting of several thousand probabilistic instances.

Besides the probabilistic set-covering instances, we also ran our code on probabilistic versions of the Single Source Capacitated Facility Location Problem (SSCFLP), Capacitated Warehouse Location Problem (CWLP) and Capacitated k-Median Problem (k-median) instances. We considered the following probabilistic version of SSCFLP.

$$\begin{aligned}
\min_{(x,y)} \quad & \sum_{i \in I, j \in J} c_{ij} x_{ij} + \sum_{i \in I} f_i y_i \\
\text{s.t} \quad & P(\sum_{i \in I} x_{ij} \geq \xi_i \ \forall j \in J) \geq p \\
& \sum_{j \in J} w_j x_{ij} \leq s_i y_i \ \forall i \in I \\
& x_{ij} \in \{0, 1\} \ \forall i \in I, \forall j \in J \\
& y_i \in \{0, 1\} \ \forall i \in I
\end{aligned} \tag{6.13}$$

Here I is the set of facilities, s_i is the capacity and f_i is the fixed cost associated with facility $i \in I$, while J is the set of customers, w_j is the demand of customer j and c_{ij} is the cost of serving customer j via the facility i for all i, j . The model obtained by replacing the integrality constraints on the x_{ij} variables in (6.13) by $0 \leq x_{ij} \leq 1$ is the probabilistic version of the CWLP. Similarly, appending the constraint $\sum_{i \in I} y_i \leq k$ ($k \geq 1$) to the probabilistic CWLP yields the probabilistic version of the k-median problem. Note that all of our results in Sections 2 (except Proposition 2.5), 3 and 5 (except Proposition 5.3) can be applied to probabilistic CWLP and k-median problems too. Since these results form the basis of the our computational experiments, we decided to include the probabilistic variants of CWLP and k-median models in our test-bed.

Table 2 gives detailed information about the deterministic instances we chose from the literature, which were subsequently used to generate a test-bed of probabilistic instances as described below. From each problem set, we retained only those instances which could be solved to optimality by the default version of CPLEX 9.0 within a time-limit of 1hr. Note that this selection criterion allows us to generate a test-bed of easy and moderately difficult instances which can be used to gain insights into the interplay between the integrality and probabilistic constraints of (PSC). Indeed, some extremely difficult set-covering and SSCFLP instances were excluded by this selection criterion; these instances are likely to give rise to

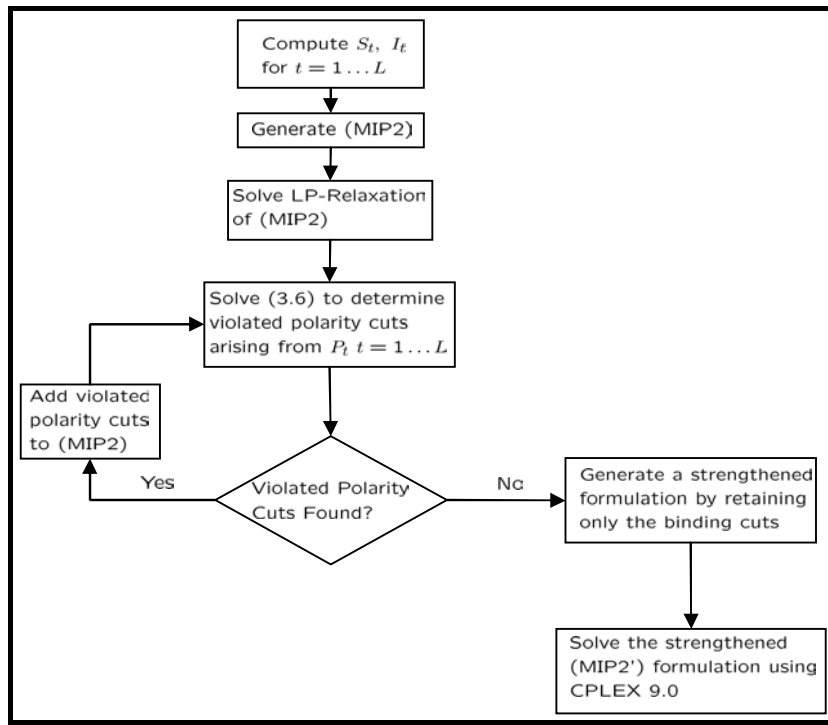


Figure 4: Flow Chart of the Algorithm

extremely difficult probabilistic instances wherein the integrality constraints of (PSC) themselves make the problem difficult to solve, their interaction with the probabilistic constraints notwithstanding.

From each deterministic instance we generated 20 probabilistic instances in the following manner. Following Beraldi and Ruszczyński [5], we considered two different block sizes, namely 5 and 10. For each one of these block sizes, we considered two different probability distributions namely, Circular and Star (see [5] for the definition of these distributions). For sake of completeness, we also considered the case of independent random variables. In particular, we have assumed that each component ξ_i ($i = 1 \dots m$) can take value 0 with probability $q_i = q_0^{1/i}$ where $0 < q_0 < 1$. Following Beraldi and Ruszczyński [5], we used $q_0 = 0.1$ in our experiments. For each one of the five combinations of block sizes and distribution type, we generated four probabilistic problems differing only in the values of the threshold probabilities which were chosen from $\{0.80, 0.85, 0.90, 0.95\}$.

For the case of Set-Covering and SSCFLP instances we used the (MIP2') formulation whereas for the case of CWLP and k-median instances we used the (MIP2) formulation. We strengthened the initial formulation by polarity cuts (section 3) for instances which were generated using the Circular and Star distribution. Since (MIP2') formulation cannot be strengthened by polarity cuts for the case of independent distribution (see section 4), the polarity cuts generator was turned off for these instances. For each probabilistic instance we ran our code with a time limit of 1hr.

Tables 10-17 summarize our key findings. The first four tables give statistics on the performance of our algorithm while the last four tables give detailed statistics on the performance of polarity cuts. The results are categorized by distribution type, block size and threshold probability p which are given in the first, second and third columns of the tables respectively. Note that for each combination of distribution type, block size and threshold probability, the test-bed had 60 set-covering instances, 70 SSCFLP instances, 37 CWLP instances and 20 k-median instances as reported in Table 2.

The fourth column in tables 10-13 gives the number of instances which could not be solved to optimality within the prescribed time-limit of 1hr. Of the 3740 probabilistic instances on which we had run our code, we were able to solve 3703 instances to optimality within 1hr. In order to assess the performance of our algorithm over the unsolved instances, we give the percentage relative gap² which remained at the end of 1hr in the fifth column of the table, averaged over instances which could not be solved to optimality. The next two columns give the total solution time and the number of branch-and-bound nodes enumerated by CPLEX, averaged over instances which could be solved to optimality within 1hr. The eighth column of the table gives the average value of the probabilistic information, calculated as $VOI = 100 \times \frac{Det(ip) - Prob(ip)}{Det(ip)}$, where VOI is the value of probabilistic information, Det(ip) is the optimal value of the deterministic problem and Prob(ip) is the optimal value of the probabilistic problem. Note that VOI represents the savings which could be made by incorporating the probabilistic information into the optimization model.

Tables 14-17 give detailed statistics on the performance of polarity cuts in our algorithmic framework. The fourth column of these tables reports the average number of rounds of polarity cuts which were generated by our code. The next two columns report the average time spent on strengthening the (MIP2) formulation by means of polarity cuts; the first column reports the total time spent on strengthening while the following column reports the time spent exclusively on solving the separation linear programs (3.6). Note that most of the time spent on strengthening was used to solve the LP relaxations of the (MIP2) formulation, and a very small fraction (less than 8% on average) was spent on solving the separation linear programs. This suggests that polarity cuts can be combined with other families of cutting planes such as Mixed Integer Gomory (MIG) cuts or Split cuts with very little computational overheads. The next two columns report the the total number of cuts which were generated and the number of cuts which were binding at the final iteration, respectively, averaged over instances in the respective category.

The next two columns report the duality gap³ closed by the (R2) and (R5) relaxations, respectively. The value reported under the (R2) column is the average duality gap closed at the root node after our code ceased to produce violated polarity cuts. The duality gap reported under the (R5) column was computed using the best solution of the (R5) relaxation

² $RG = 100 \times \frac{ip - bb}{bb}$ where RG is the percentage relative gap, ip is the value of the best solution and bb is the value of the best bound available at the end of 1hr

³ $DG = 100 \times \frac{slp - lp}{ip - lp}$, where DG is the percentage duality gap closed, lp is the value of the LP relaxation of our model, slp is the value of the relaxation and ip is the value of the optimal solution.

found by CPLEX within 1hr of computing time⁴. Note that the (R5) relaxation is significantly stronger than the (R2) relaxation for the set-covering instances, whereas for the case of SSCFLP instances both of these relaxations close almost the same duality gap. Figure 5 represents this information graphically; the vertical axis of the figure gives the ratio of the duality gap closed by (R5) and (R2) relaxations, respectively, for each value of the threshold probability p represented on the horizontal axis. Note that the strength of the (R2) and (R5) relaxations decreases as the value of the threshold probability p increases from 0.8 to 0.95, whereas the relative strength of the (R5) relaxation as compared to the (R2) relaxation increases as p increases from 0.8 to 0.95.

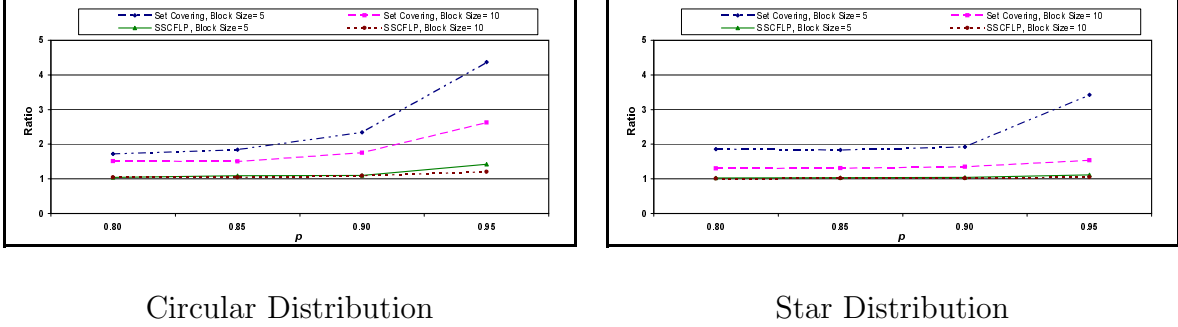
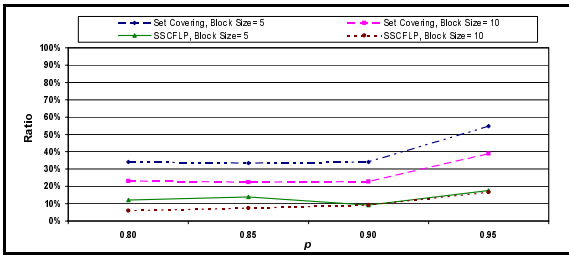


Figure 5: Ratio of Percentage Duality Gap closed by (R5) and (R2) relaxations

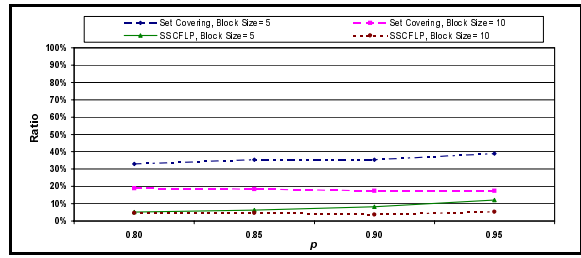
The last four columns of Tables 14-17 report statistics on the fractionality of the optimal solution to the LP relaxation of the (MIP2) formulation before and after adding the polarity cuts. Given a feasible solution (x, z, η) to the (possibly strengthened) LP relaxation of (MIP2), let $f_z = |\{i \in M \mid 0 < z_i < 1\}|$ and $f_g = |\{t \in \{1 \dots L\} \mid \exists i \in M_t \text{ s.t. } 0 < z_i < 1\}|$; f_z is a measure of fractionality of (x, z, η) in the z_i components whereas f_g measures the same in an aggregated form. The last four columns of the tables report the average values of f_z and f_g before and after adding the polarity cuts. Note that polarity cuts reduce the number of fractional z components by 80% on average. Furthermore, the impact of polarity cuts on the fractionality of the incumbent solution is more pronounced in the case of SSCFLP instances as compared to the Set-Covering instances. Figure 6 plots the ratio of the average number of fractional z components in the incumbent solution to the (MIP2) formulation before and after adding the polarity cuts.

It is interesting to note that polarity cuts do not close any fraction of the duality gap on the k-median instances. This can be attributed to the specific structure of the k-median instances in the OrLib repository. These instances were generated by choosing random points in $[0, 100] \times [0, 100]$ where every point served as a customer and potential facility, and the cost of assigning a customer to a facility is the euclidean distance between the corresponding points, rounded down to the nearest integer. Consequently, the LP relaxation of these models fractionally assigns the customer at (x, y) to the facility at (x, y) thereby giving a relaxation value of 0. The same argument carries over to the (R2) and (R5) relaxations of the probabilistic version,

⁴CPLEX was able to solve the (R5) relaxations of all the instances in our test-bed within the prescribed time-limit of 1hr for each relaxation.



Circular Distribution



Star Distribution

Figure 6: Ratio of the number of fractional z components after and before adding the polarity cuts

thus explaining the zeros in columns 9 and 10 of Table 17. Nevertheless, polarity cuts are indeed effective in decreasing the fractionality of the optimal LP relaxation solutions of these instances, as shown by the last 4 columns of Table 17.

Table 3 gives detailed information about instances which could not be solved to optimality within the prescribed time-limit of 1hr. The first four columns of the table report the problem class, distribution type, deterministic instance and block-size of the distribution, respectively. The last four columns report the values of the threshold probability p . An entry of U in the last four columns indicates that the corresponding probabilistic instance could not be solved to optimality within 1hr. It is interesting to note that the relative hardness of probabilistic instances which differ only in the value of the threshold probability p varies non-monotonically with p . For instance, our code was able to solve probabilistic variant of the SSCFLP instance $p59$ (Star distribution and block size 5) for $p = 0.80, 0.85$ and 0.90 , and yet failed to solve the instance arising from $p = 0.95$.

Next we compare our results with the earlier work of Beraldi and Ruszczyński [6] who conducted their experiments on two set-covering instances, namely $scp41$ and $scp42$, from the ORLIB repository [4]. Both of these problems have 200 set-covering constraints. They constructed 20 probabilistic instances from these two instances by considering the five combinations of group sizes and distribution types shown in Table 4, and two values of the threshold probability p , namely 0.90 and 0.95, for each combination. They tested several variants of their algorithm and concluded that a certain variant, which they refer to as the hybrid strategy with simple heuristic, performs best on their test-bed.

Table 5 compares the performance of our algorithm with the best version of the algorithm proposed in [6] on the test-bed constructed by Beraldi and Ruszczyński (see [6] for description of these instances). The first column of the table gives the problem description; a suffix of 1 (2) indicates that the instance was generated from $scp41$ ($scp42$). The second column gives the value of threshold probability p . The next two columns report the computational results of [6]; the first column gives the total computing time while the following column gives the number of p -efficient points which were enumerated by their algorithm. The next two columns report the performance of our algorithm on these instances. The first column reports the total computing time while the following column gives the number of branch-and-bound

Problem Class	Distribution	Instance	Block Size	p			
				0.80	0.85	0.90	0.95
SC	Circular	scpnre2	5	U	U		
SC	Circular	scpnre2	10	U	U	U	
SC	Circular	scpnre4	5	U			
SC	Circular	scpnrf5	5	U	U	U	
SC	Circular	scpnrf5	10	U			
SC	Star	scpnre2	5		U		
SC	Star	scpnre2	10		U	U	U
SC	Star	scpnre4	10	U			
SC	Star	scpnrf4	10	U			
SC	Star	scpnrf5	5		U		
SC	Star	scpnrf5	10		U	U	
SC	Independent	scpnre2	1				U
SC	Independent	scpnrf5	1				U
SSCFLP	Circular	p57	5			U	
SSCFLP	Circular	p57	10	U			
SSCFLP	Circular	p59	5			U	U
SSCFLP	Circular	p59	10			U	
SSCFLP	Star	p29	10	U			
SSCFLP	Star	p30	10	U			
SSCFLP	Star	p31	10	U			
SSCFLP	Star	p32	10	U			
SSCFLP	Star	p59	5				U
SSCFLP	Star	p59	10	U			
SSCFLP	Star	p71	10	U			
SSCFLP	Independent	p31	1	U			
SSCFLP	Independent	p59	1		U		
SSCFLP	Independent	p62	1	U			
SSCFLP	Independent	p70	1			U	

Table 3: Unsolved Instances

Problem	Distribution	Group Size	Number of Groups
Test 11	Star	5	40
Test 12	Circular	5	40
Test 13	Star	10	20
Test 14	Circular	10	20
Test 15	Independent	1	200

Table 4: Characteristics of Beraldi and Ruszczyński's test problems

Problem	p	BR		SGL	
		Solution Time (sec)	PEP	Solution Time (sec)	# B&B Nodes
Test 11.1	0.95	6.75	209	0.02	0
Test 11.1	0.90	60.87	935	0.03	0
Test 12.1	0.95	7.76	165	0.06	25
Test 12.1	0.90	448.52	6387	0.02	0
Test 13.1	0.95	3.99	79	0.04	5
Test 13.1	0.90	37.90	463	0.04	0
Test 14.1	0.95	8.83	252	0.02	0
Test 14.1	0.90	380.76	5284	0.02	0
Test 15.1	0.95	8780.02	140274	0.00	0
Test 15.1	0.90	33153.07	529814	0.02	0
Test 11.2	0.95	7.68	221	0.02	0
Test 11.2	0.90	123.41	1881	0.08	0
Test 12.2	0.95	4.02	106	0.06	0
Test 12.2	0.90	400.20	6249	0.04	0
Test 13.2	0.95	8.65	217	0.02	0
Test 13.2	0.90	155.64	1745	0.14	10
Test 14.2	0.95	13.32	297	0.18	27
Test 14.2	0.90	911.31	13267	0.13	6
Test 15.2	0.95	8170.48	130049	0.02	0
Test 15.2	0.90	24581.25	389886	0.02	0

Table 5: Comparison with the Beraldi and Ruszczyński Algorithm

nodes enumerated by CPLEX. Notice that the computing time of our approach is several orders of magnitude better than that of [6]. Furthermore, the extent of enumeration in our approach (column 6) is substantially smaller than in the approach proposed in [6] (column 4).

Next we discuss the importance of polarity cuts in the overall solution procedure by demonstrating their impact on solving the probabilistic version of a SSCFLP instance. We chose the instance p31 (30 facilities and 150 customers) from the Holmberg test-bed [15] and generated its probabilistic variant using the Circular distribution and threshold probability $p = 0.8$; the resulting probabilistic instance had 15 blocks of size 10 each. We ran our code on this instance in three setups. In the first setup we ran the default version of our algorithm which involves, among other things, adding polarity cuts at the root node. The second and third setups were identical to the first setup, except that the generator for polarity cuts was turned off; we used the (MIP2) and (MIP2') formulations in the second and third setups, respectively. In the first setup, our code closed 67.84% of the duality gap at the root node in less than 1sec; CPLEX acting on the formulation, strengthened by polarity cuts, was able to solve the instance to optimality in additional 52sec by enumerating 2378 branch and bound nodes. The second setup involved applying CPLEX to the unstrengthened formulation. Interestingly CPLEX, unaided by the polarity cuts, was not able to solve the instance to optimality in 2hr; it enumerated around 154,100 branch-and-bound nodes and closed only 72% of the duality gap at the end of two hours. After additional 31 hours CPLEX was able to solve the instance to optimality by enumerating 1.7 million branch-and-bound nodes. CPLEX took around 21hrs to solve the instance to optimality in the third setup and enumerated 764006 branch-and-bound nodes. Table 6 summarizes the statistics associated with these three setups. As this example demonstrates, polarity cuts have a huge impact on the overall solution time of our procedure.

% Gap closed at Root Node	67.84%
Time Spent in Strengthening	0.83 sec
Time Spent in Solving SepLP	0.30 sec
Time Taken by CPLEX 9.0 after Strengthening	52.31 sec
No. of Branch-and-Bound nodes enumerated by CPLEX 9.0	2378
Total time taken to solve the instance to optimality	53.14 sec

With Polarity Cuts (MIP2')

No. of Branch-and-Bound nodes enumerated by CPLEX 9.0	764006
Total time taken to solve the instance to optimality	77621 sec

Without Polarity Cuts (MIP2')

No. of Branch-and-Bound nodes enumerated by CPLEX 9.0	1717126
Total time taken to solve the instance to optimality	119922 sec

Without Polarity Cuts (MIP2)

Table 6: Probabilistic version of the SSCFLP instance p31

Type of Cuts	# Rounds	% Duality Gap Closed	Time (sec)
Polarity Cuts	5	67.84	0.83
CPLEX	-	48.35	16.87
RedSplit	5	8.66	4.27
	10	9.99	12.50
MIR	5	11.42	0.67
	10	12.21	1.31
MIG	5	26.07	0.40
	10	31.71	1.09
	15	34.94	1.19
	20	36.45	2.18
L&P	5	38.03	37.72
	10	44.81	95.09
	15	48.15	177.45
	20	50.01	292.46

Table 7: Comparing Polarity Cuts with other general purpose Cutting Planes

An interesting question is to determine if the strengthening which results due to addition of polarity cuts can also be obtained by adding general purpose cutting planes such as mixed integer gomory cuts or lift-and-project cuts to the (MIP2) formulation. In order to answer this question, we tried to strengthen the (MIP2) formulation of the probabilistic SSCFLP p31 instance (described above) by using other well-known classes of cutting planes. Table 7 summarizes our key findings. The first column of the table reports the type of cutting plane procedure; the second column reports the number of rounds of cuts which were generated. The third column gives the percentage duality gap closed by the respective class of cutting planes while the last column reports the total time spent on strengthening. The first row reports the performance of polarity cuts. The second row reports the performance of cuts generated by CPLEX 9.0 MIP solver at the root node; CPLEX was used in the "move best bound" mode and all of its cuts generators (except disjunctive cuts) were used in the "aggressive" mode so as to extract the best performance of the CPLEX cut generators. The remaining rows report the performance of mixed integer Gomory (MIG) cuts, mixed integer rounding (MIR) cuts, reduce-and-split (RedSplit) cuts and Lift-and-Project (L&P) cuts. We used the COIN-OR modules CglGomory, CglMixedIntegerRounding and CglRedSplit to generate MIG, MIR and reduce-and-split cuts, respectively; the lift-and-project cuts were generated using the same code as used by Balas and Saxena [2]. Note that among all classes of cutting planes, polarity cuts close the maximum fraction of the duality gap in minimum amount of time.

We conclude this section by reporting our computational experience with the (MIP4) formulation for stationary distributions. Recall that any stationary distribution is completely defined by a vector $(\lambda_0, \dots, \lambda_m)$ ($m = |M|$) where λ_i represents the value of the stationary distribution at a lattice point with exactly i ones. We used the following scheme to generate m -dimensional stationary distributions.

1. Let $a_j = \frac{100u_j}{2^{j+1}}$ $j = 0 \dots m$ where u_j is a random number in the interval $(0, 1)$.
2. Let $\mu = \sum_{j=0}^m a_j$ and let $a_j := \frac{a_j}{\mu}$ $j = 0 \dots m$.
3. For $j = 0 \dots m$, let $\lambda_j = \sum_{k=0}^j a_k \prod_{i=0}^{k-1} \frac{j-i}{m-i}$.

It can be easily verified that $(\lambda_0, \dots, \lambda_m)$ obtained by the above procedure defines a cumulative distribution function of a m -dimensional stationary distribution.

For each deterministic instance (Table 2) we generated 32 probabilistic instances in the following manner. We considered 8 different block sizes, namely 5, 10, 20, 50, $\frac{m}{4}$, $\frac{m}{3}$, $\frac{m}{2}$ and m where m denotes the number of probabilistic set covering constraints in the deterministic instance. For each block, the stationary distribution was defined using the scheme described above. For each one of these block sizes we generated four probabilistic problems differing only in the values of the threshold probabilities which were chosen from $\{0.80, 0.85, 0.90, 0.95\}$. We solved each one of the resulting probabilistic instance by CPLEX 10.1 with a time limit of 1hr.

Tables 18-21 summarizes the computational results. The columns of tables 18-21 have the same interpretation as those of tables 10-13. Our goal in this experiment was to verify whether special properties of distributions can be exploited to solve probabilistic problems with arbitrarily large block sizes. As is evident from Tables 18-21 our goal was largely attained, at least on this test-bed of problem instances.

7 Concluding Remarks

In this paper we set out to explore MIP reformulations of the probabilistic set covering problem (PSC). We formulated (PSC) as a mixed integer non-linear program (MINLP) which was subsequently linearized to obtain a MIP reformulation (MIP1). We introduced the concepts of p -inefficiency and polarity cuts. While the former was aimed at reducing the number of constraints in our model, the later was used as a strengthening device to obtain stronger formulations. A hierarchy of relaxations for (PSC) was introduced, and fundamental relationships between the relaxations were established culminating with a reformulation of (PSC) with no additional integer constrained variables. Simplifications of the MIP model which result due to special properties of matrix A and distribution function F were briefly discussed. We corroborated our theoretical findings by an extensive computational experiment on a test-bed consisting of almost 10,000 probabilistic instances. Tables 8 and 9 summarize our computational results.

Problem Class	# Probabilistic Instances	# Unsolved Instances	% RG	Solution Time (sec)	# B& B Nodes	Strengthening by Polarity Cuts	
						Time Spent (sec)	% Duality Gap Closed
Set Covering	1200	21	10.89	153.59	6434.76	0.781	31.65
SSCFLP	1400	16	0.43	36.11	2488.26	0.190	27.67
CWLP	740	0	0.00	0.35	37.65	0.046	17.17
k-Median	400	0	0.00	50.88	1636.70	0.076	0.00

Table 8: Summary Results: Circular, Star and Independent Distributions

Problem Class	# Probabilistic Instances	# Unsolved Instances	% RG	Solution Time (sec)	# B& B Nodes
Set Covering	1920	128	21.34	112.93	9373.60
SSCFLP	2240	17	0.35	9.36	1609.35
CWLP	1184	0	0.00	0.09	5.25
k-Median	640	0	0.00	2.90	156.22

Table 9: Summary Results: Stationary Distribution

This paper treads on the interface of two important areas of computational optimization - probabilistic programming and mixed integer programming. The main contribution of the paper, however, lies in integrating celebrated concepts from each one of these fields, namely p -efficiency from probabilistic programming and polarity from mixed integer programming, to create an algorithmic framework to solve (PSC) which is orders of magnitude more efficient than any of the existing approaches.

References

- [1] Balas, E.: Disjunctive Programming, Properties of the convex hull of feasible points. Discrete Applied Mathematics 89(1-3), 3-44 (1998)
- [2] Balas, E., Saxena, A.: Optimizing over the split closure. Mathematical Programming. To appear
- [3] Bartholdi, J. J.III, Orlin, J. B., Ratliff, H.D.: Cyclic scheduling via integer programs with circular ones., Operations Research 28. 1074-1085 (1980)
- [4] Beasley, J.E.: OR-Library, people.brunel.ac.uk/mastjjb/jeb/info.html
- [5] Beraldi, P., Ruszczyński, A.: The probabilistic set covering problem. Operations Research 50, 956-967 (2002)
- [6] Beraldi, P., Ruszczyński, A.: A branch and bound method for stochastic integer problems under probabilistic constraints. Optimization Methods and Software 17(3), 359-382 (2002)
- [7] Charnes, A., Cooper, W.W.: Chance constrained programming. Management Science 6, 73-79 (1959)

- [8] Christof, T., Lobel, A.: PORTA - POLYhedron Representation Transformation Algorithm, <http://www.zib.de/Optimization/Software/Porta/>
- [9] COIN: Computational infrastructure for operations research, <http://www.coin-or.org>
- [10] Cornuejols, G., Conforti, M.: Balanced 0,1,-1 Matrices, bicoloring and total dual integrality, *Mathematical Programming* 71, 249-258 (1995)
- [11] Cornuejols, G.: Combinatorial Optimization: Packing and Covering. Published by SIAM (2001) in the CBMS-NSF Regional Conference Series in Applied Mathematics CBMS 74 (2000)
- [12] Dentcheva, D., Prékopa, A., Ruszczyński, A.: Concavity and efficient points of discrete distributions in probabilistic programming. *Mathematical Programming* 89, 55-77 (2000)
- [13] Eisenbrand, F., Oriolo, G., Ventura, P., Stauffer, G.: Circular ones matrices and the stable set polytope of quasi-line graphs. *Combinatorica*. To appear
- [14] Fischetti, M., Lodi, A.: Optimizing over the first Chvatal closure. *Mathematical Programming*. To appear
- [15] Holmberg, K., Ronnqvist, M., Yuan, D.: An exact algorithm for the capacited facility location 517 problems with single sourcing. *European Journal of Operational Research* 113, 544-559 (1999)
- [16] Lejeune, M.A., Ruszczyński, A.: An efficient method trajectory method for probabilistic inventory-production-distribution problems. *Operations Research* (2007). In press
- [17] Luedtke, J., Ahmed, S., Nemhauser, G.: Strong MIP formulations for chance constrained linear programs with random right-hand side. *INFORMS Annual Meeting*, Pittsburgh (2006)
- [18] Prékopa, A.: Dual method for a one-stage stochastic programming with random rhs obeying a discrete probability distribution. *Zeitschrift of Operations Research* 34, 441-461 (1990)
- [19] Prékopa, A.: *Stochastic Programming*. Kluwer, Boston (1995)
- [20] Prékopa, A.: Probabilistic programming models. Chapter 5 in: *Stochastic Programming: Handbook in Operations Research and Management Science* 10. Edited by Ruszczyński, A., Shapiro, A. Elsevier Science Ltd, 267-351 (2003)
- [21] Prékopa, A., Vizvari, B., Badics, T.: Programming under probabilistic constraint with discrete random variable. In: *New Trends in Mathematical Programming*. Edited by Giannessi, F., Komlósi, S., Rapcsák, T. Boston, MA (1998)
- [22] de Souza, C., Balas, E.: The vertex separator problem: algorithms and computations. *Mathematical Programming* 103(3), 609-631 (2005)

Appendix

Proof of Proposition 4.2: Suppose A is a balanced matrix and $(\bar{x}, \bar{z}, \bar{\eta})$ is an optimal solution to (R5) for a given choice of c , F and p . In order to show that $\text{opt}(\text{R5}) = \text{opt}(\text{MIP2})$ it suffices to prove that $\text{opt}(\text{R5}) \geq \text{opt}(\text{MIP2})$. Consider, the following optimization problem,

$$\min \{cx \mid Ax \geq \bar{z}, 0 \leq x \leq 1\} . \quad (.14)$$

Since A is a balanced matrix every submatrix of A is ideal, which implies that (.14) has an integer optimal solution, say \hat{x} . Note that $c\hat{x} = c\bar{x}$, $(\hat{x}, \bar{z}, \bar{\eta})$ is a feasible solution to (MIP2) and hence $\text{opt}(\text{R5}) \geq \text{opt}(\text{MIP2})$.

Conversely, suppose A is not a balanced matrix. By Theorem 6.1 of Cornuejols [11], it follows that there exists a non-ideal submatrix of A . In other words, $\exists \bar{z} \in \{0, 1\}^M$ and $c \in \mathbb{R}^N$ such that the unique optimal solution to (.14) is fractional. Let $p = 0.8$ and $F : \{0, 1\}^M \rightarrow \mathbb{R}$ be defined as,

$$F(v) = \begin{cases} 1 & \text{if } v_i = 1 \ \forall i \in M \\ p & \text{if } v \geq \bar{z} \text{ and } \exists i \in M \text{ s.t } v_i = 0 \\ 0 & \text{otherwise} \end{cases}$$

for $v \in \{0, 1\}^M$. Note that \bar{z} is the unique p -efficient point of F . Consequently,

$$\text{opt}(\text{MIP2}) = \min_x \{cx \mid Ax \geq \bar{z}, x_j \in \{0, 1\} \ \forall j \in N\} \quad (.15)$$

$$> \min_x \{cx \mid Ax \geq \bar{z}, 0 \leq x \leq 1\} \quad (.16)$$

$$= \text{opt}(\text{R5}) , \quad (.17)$$

where (.16) follows from our choice of c . ■

Proof of Proposition 4.3: Let $(\bar{x}, \bar{\eta}, \bar{z}, \bar{y})$ be an optimal solution to (4.11) and let $\Delta = c\bar{x}$ be the corresponding optimal solution value. It suffices to show that $\Delta \geq \text{opt}(\text{MIP2})$. Let $Q = \{x \mid Ax \geq \bar{z}, \sum_{j \in N} x_j = \bar{y}, 0 \leq x \leq 1\}$. Following Bartholdi, Orlin and Ratliff [3] (also see Eisenbrand et al [13]) we consider the unimodular transformation $x = Tu$ where

$$T = \begin{pmatrix} 1 & & & & & \\ -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & -1 & 1 & & \\ & & & \ddots & \ddots & \\ & & & & 1 & -1 \\ & & & & -1 & 1 \end{pmatrix} .$$

The image of Q under this transformation is given by,

$$\bar{Q} = \left\{ u \in \mathbb{R}^N \mid \begin{pmatrix} A \\ -I \\ I \end{pmatrix} Tu \geq \begin{pmatrix} \bar{z} \\ -1 \\ 0 \end{pmatrix}, u_n = \bar{y} \right\} .$$

In the following we denote $\begin{pmatrix} A \\ -I \\ I \end{pmatrix} Tu \geq \begin{pmatrix} \bar{z} \\ -1 \\ 0 \end{pmatrix}$ by $Bu \geq d$. Let us write the matrix B as $B = (R|s)$ where s denotes the last column of B . Observe that, by construction, s is also

the last column of $\begin{pmatrix} A \\ -I \\ I \end{pmatrix}$. Since A has the circular ones property, each row of the matrix R has at most one entry which is $+1$ and at most one entry which is -1 . All other entries are 0. The matrix R is thus totally unimodular. Since $u_n = \bar{y}$ is an integer for $u \in \bar{Q}$, it follows that \bar{Q} is an integral polytope. Since $Q = \{x \mid T^{-1}x \in \bar{Q}\}$ and T is a unimodular matrix, Q is also an integral polytope.

Consider the following optimization problem,

$$\min \{cx \mid Ax \geq \bar{z}, \sum_{j \in N} x_j = \bar{y}, 0 \leq x \leq 1\}. \quad (.18)$$

Since Q is an integral polytope, it follows that (.18) has an integer optimal solution, say \hat{x} . Note that $c\bar{x} = c\hat{x}$, $(\hat{x}, \bar{z}, \bar{\eta})$ is a feasible solution to (MIP2), and hence $\Delta \geq \text{opt}(MIP2)$. ■

Distribution	Block Size	p	# Unsolved Instances	% RG	Solution Time (sec)	# B&B Nodes	% VOI
Circular	5	0.80	3	10.41	91.30	3736	9.66
Circular	5	0.85	2	11.81	136.67	5943	7.17
Circular	5	0.90	1	11.34	176.73	8075	4.77
Circular	5	0.95	0	0.00	152.14	8820	2.11
Circular	10	0.80	2	9.94	185.26	5067	10.17
Circular	10	0.85	1	6.63	199.52	8059	7.45
Circular	10	0.90	1	9.36	218.97	10524	5.05
Circular	10	0.95	0	0.00	170.84	9946	2.55
Star	5	0.80	0	0.00	128.08	6075	11.91
Star	5	0.85	2	9.90	81.64	4209	9.98
Star	5	0.90	0	0.00	195.51	11825	7.02
Star	5	0.95	0	0.00	133.14	8836	3.44
Star	10	0.80	2	10.96	262.00	4124	14.51
Star	10	0.85	2	16.18	208.82	3787	12.19
Star	10	0.90	2	12.13	142.89	4618	9.44
Star	10	0.95	1	4.81	178.40	9091	4.69
Independent	1	0.80	0	0.00	91.04	2829	27.55
Independent	1	0.85	0	0.00	79.86	2677	22.15
Independent	1	0.90	0	0.00	132.02	5199	16.31
Independent	1	0.95	2	11.73	108.14	4873	10.14

Table 10: Summary Results: Set Covering Instances

Distribution	Block Size	p	# Unsolved Instances	% RG	Solution Time (sec)	# B&B Nodes	% VOI
Circular	5	0.80	0	0.00	50.48	2599	8.88
Circular	5	0.85	0	0.00	29.02	1675	6.51
Circular	5	0.90	2	0.21	23.26	1779	4.21
Circular	5	0.95	1	0.85	45.21	4001	2.10
Circular	10	0.80	1	0.26	26.16	1808	7.15
Circular	10	0.85	0	0.00	72.43	2626	5.44
Circular	10	0.90	1	0.23	61.03	3574	3.40
Circular	10	0.95	0	0.00	35.79	3399	1.54
Star	5	0.80	0	0.00	30.38	1996	10.15
Star	5	0.85	0	0.00	16.22	1073	7.76
Star	5	0.90	0	0.00	11.88	963	5.36
Star	5	0.95	1	1.18	25.87	2174	2.84
Star	10	0.80	6	0.61	57.58	3054	12.64
Star	10	0.85	0	0.00	49.17	1858	9.74
Star	10	0.90	0	0.00	18.44	1028	6.73
Star	10	0.95	0	0.00	35.09	2134	3.47
Independent	1	0.80	2	0.09	58.97	6325	10.35
Independent	1	0.85	1	0.15	25.65	2497	7.90
Independent	1	0.90	1	0.01	9.84	1016	5.48
Independent	1	0.95	0	0.00	41.58	4326	2.75

Table 11: Summary Results: SSCFLP Instances

Distribution	Block Size	p	# Unsolved Instances	% RG	Solution Time (sec)	# B&B Nodes	% VOI
Circular	5	0.80	0	0.00	0.35	50	35.69
Circular	5	0.85	0	0.00	0.25	18	32.76
Circular	5	0.90	0	0.00	0.27	22	19.61
Circular	5	0.95	0	0.00	0.17	6	17.43
Circular	10	0.80	0	0.00	0.61	74	31.35
Circular	10	0.85	0	0.00	0.40	26	26.49
Circular	10	0.90	0	0.00	0.32	13	15.71
Circular	10	0.95	0	0.00	0.17	2	8.65
Star	5	0.80	0	0.00	0.49	89	32.71
Star	5	0.85	0	0.00	0.35	42	29.51
Star	5	0.90	0	0.00	0.38	48	25.84
Star	5	0.95	0	0.00	0.19	5	21.73
Star	10	0.80	0	0.00	0.90	157	15.70
Star	10	0.85	0	0.00	0.54	50	12.59
Star	10	0.90	0	0.00	0.47	62	8.78
Star	10	0.95	0	0.00	0.23	16	4.32
Independent	1	0.80	0	0.00	0.29	33	43.71
Independent	1	0.85	0	0.00	0.25	25	34.05
Independent	1	0.90	0	0.00	0.23	11	19.54
Independent	1	0.95	0	0.00	0.17	4	9.24

Table 12: Summary Results: CWLP Instances

Distribution	Block Size	p	# Unsolved Instances	% RG	Solution Time (sec)	# B&B Nodes	% VOI
Circular	5	0.80	0	0.00	55.28	1822	15.11
Circular	5	0.85	0	0.00	32.01	948	11.56
Circular	5	0.90	0	0.00	28.96	1192	7.88
Circular	5	0.95	0	0.00	19.22	1007	3.89
Circular	10	0.80	0	0.00	125.50	3864	16.67
Circular	10	0.85	0	0.00	38.12	1210	12.44
Circular	10	0.90	0	0.00	19.08	524	8.31
Circular	10	0.95	0	0.00	32.15	2189	3.52
Star	5	0.80	0	0.00	86.11	2963	21.24
Star	5	0.85	0	0.00	86.72	2669	16.96
Star	5	0.90	0	0.00	33.70	871	12.40
Star	5	0.95	0	0.00	13.89	410	6.78
Star	10	0.80	0	0.00	212.07	5850	24.64
Star	10	0.85	0	0.00	125.58	3195	19.38
Star	10	0.90	0	0.00	44.88	1558	13.70
Star	10	0.95	0	0.00	17.44	457	7.00
Independent	1	0.80	0	0.00	14.38	708	16.34
Independent	1	0.85	0	0.00	10.36	441	12.80
Independent	1	0.90	0	0.00	11.79	458	8.95
Independent	1	0.95	0	0.00	10.31	398	4.68

Table 13: Summary Results: Capacitated k-Median Instances

Distribution	Block Size	p	# Rounds	Time (sec)		# Cuts		% Gap Closed		f_g		f_z	
				Total	SepLP	Total	Binding	(R2)	(R5)	Initial	Final	Initial	Final
Circular	5	0.80	4.6	0.706	0.012	17.7	14.8	32.55	55.94	10.2	7.8	29.7	10.1
Circular	5	0.85	4.7	0.678	0.009	14.4	12.2	28.90	53.30	8.1	6.3	23.7	7.9
Circular	5	0.90	4.1	0.584	0.006	9.7	8.0	20.90	48.90	5.8	4.7	16.6	5.7
Circular	5	0.95	2.5	0.345	0.003	2.9	2.6	9.30	40.67	3.9	3.5	6.9	3.8
Circular	10	0.80	5.7	0.976	0.033	22.5	17.8	40.44	60.94	10.4	7.5	47.6	11.0
Circular	10	0.85	5.8	0.842	0.022	19.3	15.3	38.45	57.91	8.5	6.1	38.3	8.6
Circular	10	0.90	6.0	0.731	0.013	14.8	11.1	30.91	54.25	6.4	4.7	28.3	6.5
Circular	10	0.95	3.7	0.392	0.003	5.2	4.2	16.27	42.68	3.8	3.2	10.0	3.9
Star	5	0.80	4.2	0.654	0.015	19.9	15.9	31.76	59.09	12.7	9.0	36.7	12.1
Star	5	0.85	4.0	0.642	0.011	16.0	12.9	30.52	55.95	10.2	7.8	29.5	10.4
Star	5	0.90	4.2	0.597	0.009	13.4	10.7	27.47	52.70	7.8	6.1	22.4	7.9
Star	5	0.95	3.0	0.474	0.003	6.2	5.4	12.81	43.83	4.9	4.1	12.2	4.7
Star	10	0.80	6.1	1.652	0.083	29.8	23.3	53.23	69.27	14.7	9.8	84.9	16.1
Star	10	0.85	6.1	1.384	0.061	25.6	20.0	51.18	67.05	12.6	8.4	72.2	13.3
Star	10	0.90	6.4	1.094	0.034	23.2	17.6	46.62	62.85	9.8	6.5	55.4	9.7
Star	10	0.95	5.7	0.748	0.013	16.5	12.1	35.08	53.74	6.3	4.4	34.0	5.9

Table 14: Performance of Polarity Cuts: Set Covering Instances

Distribution	Block Size	p	# Rounds	Time (sec)		# Cuts		% Gap Closed		f_g		f_z	
				Total	SepLP	Total	Binding	(R2)	(R5)	Initial	Final	Initial	Final
Circular	5	0.80	7.6	0.123	0.011	21.9	21.0	30.71	32.08	6.5	1.0	27.5	3.3
Circular	5	0.85	7.7	0.121	0.009	20.2	18.6	27.79	30.30	5.3	1.1	23.4	3.2
Circular	5	0.90	10.0	0.135	0.009	21.4	17.7	23.04	25.39	4.3	1.0	19.3	1.8
Circular	5	0.95	3.5	0.046	0.002	4.8	4.2	4.66	6.64	2.8	1.0	6.1	1.1
Circular	10	0.80	9.6	0.218	0.053	22.6	19.5	33.16	34.54	6.1	1.0	28.8	1.7
Circular	10	0.85	10.2	0.171	0.024	24.7	20.0	29.59	30.75	5.4	1.1	25.3	1.9
Circular	10	0.90	9.9	0.140	0.015	23.0	16.1	23.61	25.68	4.1	1.1	19.8	1.8
Circular	10	0.95	3.8	0.052	0.003	5.1	4.2	6.65	7.99	2.3	1.0	6.4	1.1
Star	5	0.80	5.9	0.106	0.010	18.4	17.7	31.83	32.58	8.8	1.1	31.1	1.6
Star	5	0.85	6.4	0.102	0.009	17.9	17.0	29.15	30.18	7.1	1.1	25.6	1.6
Star	5	0.90	7.8	0.108	0.008	17.9	16.1	24.07	25.12	5.5	1.0	19.6	1.6
Star	5	0.95	6.2	0.078	0.004	11.1	10.0	12.43	13.89	4.0	1.1	11.3	1.4
Star	10	0.80	17.8	0.589	0.128	42.6	27.5	49.29	49.95	7.7	1.1	52.0	2.4
Star	10	0.85	20.4	0.550	0.113	54.5	33.6	46.92	47.78	6.8	1.1	46.5	2.2
Star	10	0.90	18.7	0.346	0.057	48.7	30.1	41.72	42.58	6.0	1.1	41.7	1.5
Star	10	0.95	12.9	0.161	0.018	30.8	20.1	28.07	29.41	4.4	1.1	27.2	1.4

Table 15: Performance of Polarity Cuts: SSCFLP Instances

Distribution	Block Size	p	# Rounds	Time (sec)		# Cuts		% Gap Closed		f_g		f_z	
				Total	SepLP	Total	Binding	(R2)	(R5)	Initial	Final	Initial	Final
Circular	5	0.80	5.0	0.028	0.002	5.5	5.2	6.35	10.82	2.5	1.0	5.6	2.5
Circular	5	0.85	3.6	0.025	0.001	3.4	3.2	3.79	11.06	2.0	1.0	3.7	1.3
Circular	5	0.90	2.0	0.014	0.001	1.1	1.1	1.28	39.96	1.1	1.0	2.5	1.0
Circular	5	0.95	1.1	0.009	0.000	0.1	0.1	0.06	1.13	1.1	1.0	1.3	1.0
Circular	10	0.80	7.8	0.111	0.031	17.6	15.1	33.78	37.10	4.0	1.1	22.8	1.6
Circular	10	0.85	7.5	0.071	0.012	12.3	10.6	29.34	35.26	3.2	1.0	18.1	1.5
Circular	10	0.90	7.5	0.051	0.006	10.7	7.6	14.41	35.28	2.2	1.0	13.5	1.0
Circular	10	0.95	2.4	0.015	0.001	1.4	1.4	1.20	5.44	1.1	1.0	2.1	1.0
Star	5	0.80	4.5	0.033	0.004	8.1	7.5	21.99	37.81	4.7	1.1	15.4	1.2
Star	5	0.85	4.2	0.035	0.004	7.0	6.6	24.57	26.66	4.7	1.0	14.3	1.4
Star	5	0.90	6.6	0.045	0.005	10.8	9.1	24.49	27.35	3.9	1.1	12.3	1.8
Star	5	0.95	2.9	0.021	0.002	4.0	3.7	5.73	9.22	3.0	1.0	6.4	1.1
Star	10	0.80	9.7	0.102	0.020	15.0	12.7	38.15	40.42	4.1	1.1	24.0	1.8
Star	10	0.85	11.5	0.096	0.015	18.7	13.2	35.02	37.69	4.1	1.2	22.9	2.9
Star	10	0.90	5.8	0.046	0.008	10.3	8.0	25.70	28.77	3.9	1.1	18.7	2.0
Star	10	0.95	6.1	0.039	0.004	8.1	6.7	8.81	11.47	2.1	1.0	10.4	1.0

Table 16: Performance of Polarity Cuts: CWLP Instances

Distribution	Block Size	p	# Rounds	Time (sec)		# Cuts		% Gap Closed		f_g		f_z	
				Total	SepLP	Total	Binding	(R2)	(R5)	Initial	Final	Initial	Final
Circular	5	0.80	2.6	0.058	0.002	4.6	4.5	0.00	0.00	4.1	0.7	6.8	0.8
Circular	5	0.85	2.5	0.055	0.001	3.7	3.7	0.00	0.00	3.3	0.6	5.8	0.6
Circular	5	0.90	2.5	0.053	0.001	2.8	2.7	0.00	0.00	2.6	0.9	4.3	1.1
Circular	5	0.95	2.2	0.036	0.000	1.5	1.5	0.00	0.00	1.4	0.5	2.4	0.5
Circular	10	0.80	3.1	0.103	0.010	6.8	6.5	0.00	0.00	4.7	0.5	20.6	0.5
Circular	10	0.85	4.0	0.097	0.007	6.7	5.8	0.00	0.00	3.8	0.7	17.2	0.8
Circular	10	0.90	4.5	0.075	0.004	6.0	4.2	0.00	0.00	2.8	0.9	10.3	1.0
Circular	10	0.95	2.1	0.037	0.000	1.1	1.0	0.00	0.00	1.1	0.5	2.6	0.5
Star	5	0.80	2.3	0.062	0.003	6.3	6.1	0.00	0.00	6.1	1.0	13.5	1.0
Star	5	0.85	2.1	0.055	0.002	5.5	5.5	0.00	0.00	5.5	0.9	11.7	0.9
Star	5	0.90	2.4	0.056	0.001	4.9	4.7	0.00	0.00	4.4	0.6	9.6	0.8
Star	5	0.95	2.4	0.045	0.001	2.5	2.4	0.00	0.00	2.3	0.7	4.2	0.7
Star	10	0.80	2.4	0.151	0.031	5.4	5.1	0.00	0.00	5.0	1.0	28.7	1.1
Star	10	0.85	3.3	0.141	0.024	6.2	5.1	0.00	0.00	4.6	1.0	27.3	1.0
Star	10	0.90	4.2	0.112	0.012	7.3	5.4	0.00	0.00	4.0	0.8	24.0	0.8
Star	10	0.95	4.8	0.084	0.005	6.4	4.6	0.00	0.00	2.8	0.9	11.8	1.1

Table 17: Performance of Polarity Cuts: Capacitated k-Median Instances

Block Size	p	# Unsolved Instances	% RG	Solution Time (sec)	# B&B Nodes	% VOI
5	0.80	3	13.74	120.68	7732	6.09
5	0.85	0	0.00	223.13	18120	4.35
5	0.90	1	5.10	86.89	8790	2.68
5	0.95	0	0.00	84.24	10482	0.18
10	0.80	7	14.98	81.12	4445	8.86
10	0.85	4	13.64	132.8	8941	5.91
10	0.90	0	0.00	117.88	11263	3.77
10	0.95	0	0.00	74.81	10558	0.12
20	0.80	10	20.53	28.96	2264	13.13
20	0.85	6	18.91	238.04	9811	10.14
20	0.90	4	9.72	146.56	9543	6.12
20	0.95	0	0.00	56.09	8346	3.07
50	0.80	10	34.14	77.56	6695	20.15
50	0.85	10	24.58	22.55	2561	16.38
50	0.90	10	13.28	6.85	794	10.46
50	0.95	0	0.00	124.86	13138	4.84
$m/4$	0.80	11	29.00	102.37	6415	26.80
$m/4$	0.85	11	23.76	98.6	7544	20.10
$m/4$	0.90	7	22.80	113.66	11077	13.41
$m/4$	0.95	2	19.88	129.12	12252	6.29
$m/3$	0.80	9	23.92	76.13	6197	30.25
$m/3$	0.85	8	19.78	71.84	5024	24.11
$m/3$	0.90	5	17.87	232.51	16817	16.21
$m/3$	0.95	1	10.30	174.66	19668	8.02
$m/2$	0.80	4	22.40	120.51	6492	39.51
$m/2$	0.85	2	19.26	159.57	9484	31.94
$m/2$	0.90	2	22.74	90.4	6859	23.08
$m/2$	0.95	0	0.00	155.09	13958	12.15
m	0.80	0	0.00	70.03	5678	46.23
m	0.85	0	0.00	130.65	11764	34.40
m	0.90	0	0.00	139.26	11943	23.93
m	0.95	0	0.00	83.51	9343	12.23

Table 18: Summary Results for Stationary Distribution: Set Covering Instances

Block Size	p	# Unsolved Instances	% RG	Solution Time (sec)	# B&B Nodes	% VOI
5	0.80	0	0.00	9.79	1177	5.86
5	0.85	0	0.00	7.09	1547	3.96
5	0.90	0	0.00	4.78	1066	2.24
5	0.95	1	0.40	4.4	884	0.14
10	0.80	2	0.16	45.14	7281	7.12
10	0.85	2	0.45	12.09	1822	4.68
10	0.90	1	0.53	5.46	1185	2.70
10	0.95	0	0.00	3.06	671	0.00
20	0.80	1	1.39	9.17	1131	11.25
20	0.85	0	0.00	3.61	551	7.68
20	0.90	0	0.00	10.01	1720	3.86
20	0.95	0	0.00	6.25	1128	0.55
50	0.80	1	0.77	2.7	110	34.73
50	0.85	0	0.00	30.3	3395	21.02
50	0.90	0	0.00	2.17	198	10.44
50	0.95	2	0.06	23.53	4796	3.88
$m/4$	0.80	0	0.00	8.11	1540	11.81
$m/4$	0.85	0	0.00	12.94	1898	7.92
$m/4$	0.90	0	0.00	2.81	313	4.12
$m/4$	0.95	1	0.74	5.57	1096	1.04
$m/3$	0.80	1	0.26	7.5	1113	17.39
$m/3$	0.85	0	0.00	2.88	320	10.65
$m/3$	0.90	0	0.00	6.32	1590	5.43
$m/3$	0.95	0	0.00	6.54	1425	1.51
$m/2$	0.80	0	0.00	2.09	103	29.62
$m/2$	0.85	0	0.00	6.95	1316	17.97
$m/2$	0.90	3	0.07	37.43	8678	10.02
$m/2$	0.95	0	0.00	3.94	730	3.77
m	0.80	0	0.00	12.1	1965	37.17
m	0.85	2	0.17	3.89	871	23.45
m	0.90	0	0.00	1.29	131	13.43
m	0.95	0	0.00	1.76	230	5.00

Table 19: Summary Results for Stationary Distribution: SSCFLP Instances

Block Size	p	# Unsolved Instances	% RG	Solution Time (sec)	# B&B Nodes	% VOI
5	0.80	0	0.00	0.10	5	20.96
5	0.85	0	0.00	0.08	3	19.53
5	0.90	0	0.00	0.07	2	2.78
5	0.95	0	0.00	0.07	2	0.00
10	0.80	0	0.00	0.24	28	38.16
10	0.85	0	0.00	0.14	15	29.65
10	0.90	0	0.00	0.08	4	19.70
10	0.95	0	0.00	0.07	2	3.77
20	0.80	0	0.00	0.19	14	51.39
20	0.85	0	0.00	0.12	11	44.49
20	0.90	0	0.00	0.08	2	34.36
20	0.95	0	0.00	0.07	2	15.00
50	0.80	0	0.00	0.08	3	45.45
50	0.85	0	0.00	0.07	3	30.40
50	0.90	0	0.00	0.07	3	19.54
50	0.95	0	0.00	0.07	2	0.00
m/4	0.80	0	0.00	0.10	4	33.60
m/4	0.85	0	0.00	0.09	8	30.01
m/4	0.90	0	0.00	0.07	1	24.68
m/4	0.95	0	0.00	0.07	2	0.00
m/3	0.80	0	0.00	0.12	11	36.15
m/3	0.85	0	0.00	0.11	10	21.14
m/3	0.90	0	0.00	0.07	2	14.50
m/3	0.95	0	0.00	0.07	2	0.00
m/2	0.80	0	0.00	0.12	8	63.13
m/2	0.85	0	0.00	0.09	3	54.65
m/2	0.90	0	0.00	0.07	3	42.18
m/2	0.95	0	0.00	0.07	3	17.39
m	0.80	0	0.00	0.07	0	64.13
m	0.85	0	0.00	0.08	4	53.63
m	0.90	0	0.00	0.07	3	39.61
m	0.95	0	0.00	0.07	3	16.67

Table 20: Summary Results for Stationary Distribution: CWLP Instances

Block Size	p	# Unsolved Instances	% RG	Solution Time (sec)	# B&B Nodes	% VOI
5	0.80	0	0.00	4.29	113	13.95
5	0.85	0	0.00	3.52	102	10.44
5	0.90	0	0.00	2.38	93	6.50
5	0.95	0	0.00	6.16	638	0.00
10	0.80	0	0.00	4.56	174	15.31
10	0.85	0	0.00	3.95	210	10.57
10	0.90	0	0.00	2.23	109	8.47
10	0.95	0	0.00	4.81	451	1.14
20	0.80	0	0.00	2.96	74	16.74
20	0.85	0	0.00	1.95	53	10.51
20	0.90	0	0.00	2.42	136	5.68
20	0.95	0	0.00	5.89	609	0.29
50	0.80	0	0.00	2.31	180	60.50
50	0.85	0	0.00	1.76	33	43.22
50	0.90	0	0.00	1.98	74	23.41
50	0.95	0	0.00	3.67	266	7.23
m/4	0.80	0	0.00	4.24	142	23.47
m/4	0.85	0	0.00	2.96	92	15.83
m/4	0.90	0	0.00	2.60	103	7.56
m/4	0.95	0	0.00	4.95	430	1.51
m/3	0.80	0	0.00	3.69	113	22.80
m/3	0.85	0	0.00	2.22	46	15.75
m/3	0.90	0	0.00	1.50	43	8.20
m/3	0.95	0	0.00	2.45	164	2.04
m/2	0.80	0	0.00	1.22	31	39.63
m/2	0.85	0	0.00	1.34	28	28.95
m/2	0.90	0	0.00	1.45	27	16.36
m/2	0.95	0	0.00	3.15	213	5.11
m	0.80	0	0.00	0.84	11	44.21
m	0.85	0	0.00	0.97	9	29.43
m	0.90	0	0.00	1.53	44	17.06
m	0.95	0	0.00	2.90	188	6.56

Table 21: Summary Results for Stationary Distribution: Capacitated k-median Instances